

How to shoot down spaceships

Mr Caspar Weinberger, the US Secretary of Defense, seems to have put paid to international participation in the US space station. Some putative partners may be grateful to him.

It is always refreshing when powerful statesmen are willing to make their opinions public. Mr Michael Foot, when a minister in the British government and as leader of the Labour Party opposition in the House of Commons, was a frequent contributor to the newspaper correspondence columns. Mr Caspar Weinberger, the US Secretary of Defense, is similarly unafraid to let what he thinks become generally known. Even those who may believe that his article in the *New York Times* ten days ago about the prospects for arms control was a tactless public warning to his colleague Mr George Schultz, the US Secretary of State, will have been glad of the information it contains. Almost at the same time, it now appears, Weinberger was writing privately to Schultz to warn him against an over-obtrusive foreign presence on the US space station planned for 1994 (see *Nature* 326, 628; 1987). In the way that these things happen, Weinberger's letter seems to have enjoyed almost as wide and rapid a circulation as if that, too, had been sent to a newspaper. But, apart possibly from Schultz, the most attentive readers of the letter will be the putative partners in the venture — Canada, the members of the European Space Agency (ESA) and Japan. They may not be as downcast by its tone as may be thought.

There is nothing new or surprising in the US military's declaration of an interest in the project. So much has been made plain from the outset. It would be surprising if it were otherwise; there are a great many ways in which military people could use a space station without compromising the US government's understanding with its potential partners that the enterprise was intended to be essentially civilian. The military, who have forced the pace of observation in many fields of geophysics, oceanography for example, no doubt have a great deal of data-gathering to do in relatively low orbits about the Earth. Moreover, it would be disingenuous to pretend that everything done in such a laboratory should be open to public inspection; indeed, among the potential partners are some who believe there is a commercial fortune to be made from microgravity experiments, and who plan that some of their nationals should carry out experiments whose results will be kept secret even from the other occupants of the rig, and which may never be published. The notion that there might also be military programmes of work seems to have been more or less accepted until late last year, when the Pentagon asked the National Aeronautics and Space Administration (NASA), the lead agency for the project, to match its developing contractual relationships with its partners overseas against the possible interests of the US military.

Unanticipated needs

Things have not gone right since then. There is no reason to dispute the Pentagon's flat assertion that it has nothing particular in mind, but that it is merely concerned that unanticipated future needs will not be denied by arrangements made now. But the military might have guessed that asking for a formal review last December would have confirmed some of the partners in their indecision about the wisdom of participation. It is an open secret that some of the foreign candidates

for partnership plan to participate only because others are going along, and because they may be left behind. Canada is probably in this case, together with several ESA members. Other governments, that of the Netherlands for example, are genuinely alarmed at the political capital their opponents would make of what might be presented as tainted collaboration. And what is to be made of the British government, which has set up its own National Space Centre but has (so far) been unable to find the budget to make the centre a reality? That is hardly an augury of enthusiastic participation. The plain truth is that the more insistently the Pentagon has asked that its unperceived needs should be accommodated, the more zealous some partners have become in demanding that everything should be spelled out in advance. Cynics might be forgiven for thinking that some of the difficulties raised have been meant to dash the prospects for collaboration altogether.

Strong words

Now, Weinberger's letter probably ensures that that is what will happen. It is the tone that will do the damage. Literally, it asks that the United States should "resist compromise" with its partners on four stated principles — the right to conduct "national security activities" on the US elements of the structure without approval or review; there must be no "multilateral decision-making" on the management, operation and use of the space station; no "one-way flow of US space technology" (*sic*); and "equal partnership" must not displace either "the reality or the symbol of US leadership" in the project. Like the quarrelsome partners, Weinberger is simply asking that the fine print should be made plain, and that the minor partners should be told of the degree of influence to which their small contributions will entitle them. The obvious difficulty, for the putative partners, is that the message will make them more and not less intransigent.

In the long run, none of this will matter very much. Even now, there is a chance that what looks like the beginning of a substantial transatlantic row will be papered over. But there is a sense in which the gilt, such as it ever was, has been rubbed off the space station before any metal has been cut. Where Europe (represented by ESA) is concerned, there is anyway a strong case for a reconsideration of what needs to be done, in fields of technology of which space, and man in space in particular, is only one. If cool consideration showed that the first need is to make the Ariane rocket work properly, and better, that would be no great surprise. Beyond that are a string of smaller but still costly projects that need to be carried through. Japan is in much the same case. If Canada feels too neglected, it might join one or the other. Some decades hence, the circumstances may be very different. Meanwhile, the most urgent need is that the tradition of collaboration on the exploration of the Solar System between the United States and the partners in the space station and also the Soviet Union should not be allowed to lapse. If NASA is likely to be preoccupied with the space shuttle and the space station for at least the next six years, might not the minor partners more usefully devote themselves to collaborative exploration? □



Where next the dollar?

The calamitous slump of the US dollar is no more welcome because expected.

THOSE who now wonder why the US dollar should have lost 40 per cent of its value against the Japanese yen since the beginning of the year should have been reading *Nature* or a reputable economic journal all the time. The underlying causes are the federal government's budget deficit, likely to be \$150,000 million this year (whatever Gramm-Rudman says it should be next year), and the US population's belief that it has better things to do with its after-tax income than lend it to the government (or the other institutions, such as the financial institutions, from which the federal government borrows). Hitherto, the mismatch and the instability that it entails has not mattered very much. People elsewhere, corporate and private, have been willing to bridge the gap. The Japanese have had enough surplus cash both to drive the Tokyo Stock Exchange to unknown peaks and to snap up assets in the United States, making paper fortunes for their owners which then find their way to the US Treasury. Those oil-producers still in surplus have followed suit. The arithmetical result has been an imbalance in the external trade accounts of the United States.

Many in the US government appear to have believed that this genteel process could continue indefinitely, although that was always an absurdity; there must always have been a numerical upper limit to the net worth of the United States. Since the beginning of 1986, the markets have been correcting the twin imbalances in the only way they know, by depreciating the dollar and appreciating the currencies of the surplus countries. It is a remarkable testimonial to the productiveness of the US economy that this trend should not sooner have dried up the flow of investment into the United States; people do not ordinarily like buying assets whose value promptly falls.

What seems this year to have broken the spell is the administration's budget for the financial year whose beginning is still half a year away, and which seemed to many to be as artificial a way of meeting the interim Gramm-Rudman target deficit as that the US Congress is likely to adopt in a pre-election year. The final straw has been the wild talk in Washington about sanctions against Japan. The \$300-million-worth of goods on which penal tariffs may eventually be levied are numerically irrelevant; what has stuck in the craw of market-traders elsewhere is the demonstration that the United States does not understand the dilemma in which it has placed itself—that one cannot run a budget deficit and a low savings ratio without a matching trade deficit.

The danger now, not just for the United States, is multivalent. Domestically, inflation and interest rates may rise, further slowing industrial activity and slicing substantial portions off recent paper fortunes. Second, the volume of world trade may precipitously decline, with the consequence that countries in poor shape already will be even worse off. Third, and worst of all, these short-term changes could precipitate the most threatening instability of all — the huge indebtedness (exceeding \$500,000 million) of the world's debtor nations to the world's commercial banks. People worry when the engine of the free market, the US economy, falters by a per cent or so. What would be the deflationary consequences if paper assets equivalent to 10 per cent of its annual production suddenly proved worthless? Whatever happens, there will be a lot to learn from the turbulence of the past few months; economists (and national treasurers) may be persuaded that, if they want to control the international economy, they had better seek to control rates of change than numerical parities.

Mercifully, there are more immediate lessons to be learned, many of them in Washington. If the key to the present

instability is the budget deficit, it must surely help to keep the deficit, now and in the future, as small as possible. The administration has had to use some sleight of hand to keep the projected deficit for next year within the Gramm-Rudman limit of \$140,000 million, but the Congress overrode (last month) a presidential veto on its Highway Bill that will make the target even less attainable. Yet, curiously, this happening has been presented as a political defeat for the White House rather than as a serious setback for the US economy. It is as if the administration, which professes its belief that the deficit is important, does not believe that the numbers are significant. How else, for example, could Mr Caspar Weinberger, the US Secretary of Defense, have last week written a letter to the Secretary of State (see *Nature* 326, 727; 1987 and 326, 627–628; 1987) implying that a prospective foreign contribution of 15 per cent to the proposed US space station, worth perhaps \$2,000 million, counts for nothing against the encumbrance of having strings attached? No official of the Japanese government (which is perpetually in hock to its would-be pensioners) would be allowed to behave in such a cavalier fashion. Is that another lesson in competitiveness to be learned? □

Mentally speaking

The British Royal Family has an opportunity to foster openness on psychiatric illness.

THE British royal family has been embarrassed, in the past few weeks, by the disclosure (originally in *The Times*) that three cousins of the present Queen have been confined in a psychiatric hospital for more than three decades. Part of the difficulty is that the need for such treatment should have arisen in the first place. Another is that the unfortunate people affected should have been confined (admittedly when that was the standard method of treatment), but then kept confined when the ill-effects of institutionalization were recognized about a quarter of a century ago. It may also be mildly embarrassing, to a family whose genealogy must be better known than even that of Abraham's family, that these distressing circumstances had hitherto escaped attention.

In the monarchies of Western Europe, the guiding principle seems to be that royal families typify and, by so doing, dignify what may be called ordinary families. If anything, the British royal family has done a little better than might have been expected. Confining the psychiatrically sick and the mentally defective was common just a few decades ago. More recently, the Prince of Wales, the monarch in waiting, has taken an active part in several good causes of this kind. He has lent his name to an appeal for funds for schizophrenia research. Last September, at the Harvard anniversary celebrations, he made an eloquent plea for a better understanding of the "dark side" of human nature. May he, and his relatives, now be emboldened to go a little further?

One of the startling anomalies of the present time is the contrast between the persistence of reticence about psychiatric illness and the explicitness with which the problem of AIDS (acquired immune deficiency syndrome) is discussed. The explanation is not difficult. Although AIDS is mostly a venereal disease, and thus shaming, there is also a danger that there will be an epidemic and thus an urgent need for public education, with the result that reticence is overridden. Psychiatric illness is also shaming, partly because of the general belief that it "runs in families"; itself supported by observations such as that a few weeks ago that, among the Amish of southern Pennsylvania, the propensity to manic-depressive illness is genetically determined. The irony is that the concealment of most families' shame contributes directly to their individual sense of having to carry a secret burden. When he has a chance, the Prince of Wales might do a power of good by making a few kind remarks about his forgotten relatives. □

US cancer treatment gains called into question

- National Cancer Institute criticized
- Survival improvements overstated

Washington

A CONGRESSIONAL report claiming that increases in the US cancer survival rate over the past 30 years have been exaggerated has angered officials of the US National Cancer Institute (NCI). The report*, from the General Accounting Office (GAO), says that progress in extending the lives of cancer patients has been "limited" except in certain rarer cancers, and that improvement in survival for specific cancers is "often not as great as that reported". Scientists at NCI believe, however, that it is possible to draw quite different conclusions from the report.

The report was prepared at the request of Representative Ted Weiss (Democrat, New York) who was concerned about criticism he had heard of NCI figures given to Congress and to the press. Among past NCI claims that had raised eyebrows in the scientific community was one that cancer deaths could be cut in half by the year 2000 (see *Nature* 324,9; 1986). To compile the report, data on disease trends were collected from NCI documents and from an extensive literature search. Expert testimony was then taken for each of the 12 cancers studied from group interviews at national cancer centres. The NCI figures were not themselves questioned, but their interpretation certainly was.

From the expert testimony came evidence that cancer survival rates may be subject to measurement bias. Cancers are being detected earlier; changes have taken place in classifying the stage to which a cancer has developed and in what is counted as cancer; and those attending screening programmes are healthier and better educated. All these biases help to boost figures showing that people survive cancer longer. Looking at the NCI figures in this light led Weiss to the harsh conclusion that "NCI has apparently overstated the degree of genuine progress made. . . . Neither congressional policymakers nor the public is well-served by unwarranted expectations that we have turned the corner on this group of devastating diseases."

Scientists at NCI, on the other hand, read the report in a much more positive light. Great strides are documented in treating leukaemias and lymphomas. Five-year survival rates have increased over the past thirty years from 10 per cent to 33 per cent for leukaemias and

from 31 per cent to 48 per cent for non-Hodgkin's lymphoma. Carcinomas have proved more intractable but some progress has still been made. After considering possible bias the report concludes that there has been moderate improvement in survival for bladder, endometrial and prostate cancers, and slight improvement for breast, cervical, colorectal, head and neck, and small-cell carcinomas of the lung. Only for stomach cancer has there been no improvement. There have also been substantial improvements in the quality of life for cancer patients, Dr Gregory Curt, an NCI medical oncologist stressed, and these are acknowledged in the report. New diagnostic techniques now mean that investigative surgery is often unnecessary, for example. And mastectomy operations are now performed far less frequently than 20 years ago.

The biases affecting cancer survival figures do not come as news to NCI scientists; the difficulties in interpreting epidemiological data are well known. Although the Department of Health and Human Services, responsible for the NCI, has accepted the report's recommendation that survival rate data should in future include a description of possible biases, this is not a major policy change. According to Curt, NCI technical reports already include such cautions.

Whether the real gains in treating cancer have been overplayed by NCI must remain partly a matter of opinion. The report acknowledges as much—but points out that independent experts agreed with the essence of the report. Ambiguity also remains over how to rate overall improvements in cancer therapy, largely because the biggest gains have been seen in the rarest cancers: the report limply concludes that "it is difficult to find there has been much progress, but it is also impossible to say that there has been none".

Whichever way it is read, the report is not going to make it easier for NCI to gain research funds. And as 65 per cent of funds supports basic research, studies on the molecular biology of cancer could be adversely affected. The enormous progress made in the past few years has yet, of course, to have an impact on cancer survival rates. **Alun Anderson**

*Cancer patient survival: What progress has been made? GAO/PEMD-87-13

Animals can now be patented

Washington

THE US Patent and Trademark Office will announce this week that it will now consider patent applications for "nonnaturally occurring non-human multicellular living organisms" — that is, genetically altered higher animals. This policy turnaround results from a case brought by the University of Washington over a genetically manipulated oyster.

Earlier this month, the Board of Patent Appeals and Interferences overturned a rejection of an application for a Pacific oyster made sterile. The original examiner of the University of Washington application refused to grant a patent for the oyster on the basis that it was a higher life form, and therefore was not patentable subject matter.

But the appeals board ruled that that was not sufficient grounds to reject a patent application, given the Supreme Court ruling in 1980 in the *Diamond v. Chakrabarty* case that cleared the way for the patenting of microbes. According to the board, it was the intent of the law that "everything under the Sun made by man" be eligible for patent protection.

To receive a patent, the University of Washington must still prove that its sterile oyster, edible year-round, is more than an obvious extension of similar work on American oysters published previously.

This case has set the stage for the world's first patented animal. According to Rene Tegtmeier, the Assistant Commissioner of the US Patent and Trademark Office, there are 15 applications for animal patents waiting in the wings.

Humans were expressly excluded from patent eligibility in the patent board's decision. Donald J. Quigg, Assistant Secretary of Commerce and Commissioner of Patents and Trademarks, indicated that "the grant of a limited, but exclusive property right in a human being is prohibited by the Constitution".

Hybrids and cross-breeds arising from age-old breeding techniques will not be eligible for patents under the recent decision either. Applicants for animal patents will be forced to prove that the animal constitutes a manufacture or composition of matter not found in nature.

Critics of the decision state that granting patents for higher life forms implies that man can create and claim ownership for new living things. Jeremy Rifkin of the Foundation on Economic Trends has formed a coalition to attempt to have the patent office's decision rescinded. Rifkin has drafted legislation that would bar the patenting of vertebrates and invertebrates, and is now seeking a member of Congress to sponsor the bill. Carol Ezzell

US, Soviets resume space collaboration

Washington

AFTER a lapse of five years, the United States and the Soviet Union will soon officially join forces on a number of space and astronomy projects, including the exploration of Mars. On 15 April, in Moscow, Secretary of State George Shultz and Foreign Minister Edvard Shevardnadze put their names to a wide-ranging treaty specifying 16 areas in which joint efforts are proposed. The previous US-Soviet collaboration, of which the Apollo-Soyuz mission was the centrepiece, was cancelled in 1982 after the imposition of martial law in Poland.

The treaty was negotiated in October last year, but the signing had been delayed until a sufficiently high-level meeting. In the meantime, interagency bargaining in the US reduced a longer list of collaborative projects down to the sixteen now agreed upon. Certain topics involving advanced technology were omitted, and for the projects that remain, there will be no action until a review process ensures that nothing "unwarranted" is revealed to the opposite side.

Both sides are looking forward to a resumption of some plans and a beginning to others. Peter Smith, of the Division of International Affairs at the National Aeronautics and Space Administration (NASA), says that scientists will be trying to "pick up the pieces" after a five-year halt.

The new agreement covers a lot of ground. In space exploration, both Mars and Venus are immediate targets. The Mars Observer programme for remote sensing of the Martian surface will be joined with a Soviet plan to land an unmanned vehicle on Phobos, the larger of Mars's two moons, and the US Deep Space Network will be used to track the lander's movements. For Venus, radar data from the Venera 15 and 16 probes will help the United States to plan their Magellan mission, data from which will then be shared.

Informal contacts between individual scientists have proceeded, with ups and downs, in the absence of official sanction, but the new treaty, according to Smith, will make possible co-ordinated efforts in large, long-term projects.

A particularly exciting prospect for astronomers involves the Soviet plan to put a radiotelescope into orbit. Linked with other radiotelescopes around the world, it would make possible interferometry with a baseline of enormous length, allowing the structure of quasars, for instance, to be resolved in fine detail.

David Lindley

Stanford Linear Collider set for highest energies

Palo Alto, California

THE Stanford Linear Accelerator Center last week proudly showed off the new Stanford Linear Collider (SLC), expected by summer to produce the world's most powerful electron-positron collisions.

To achieve SLC's collision energies of 100 GeV, Stanford spent \$115 million improving the 21-year old Stanford Lin-



Z⁰ factory takes shape

ear Accelerator. Two hundred of the 244 original XK5 klystrons, designed to accelerate electrons and positrons to 20 GeV, have been replaced by the more powerful 5045 klystrons. After bunching to improve luminosity and damping to reduce transverse emittance, positrons and electrons are injected into the two-mile-long main accelerator. When the electron and positron bunches reach 50 GeV they are sent down two arcs, bringing them to the collision area in the large detector hall.

Construction of SLC was completed last month, and the system is now being tuned for studying the Z⁰, one of the

particles essential to the Weinberg-Salam theory of electroweak unification. The chief advantage of linear accelerators over their circular counterparts is that they do not suffer energy losses from synchrotron radiation, a consequence of beam bending. The disadvantage of the linear machines, however, is that the beam must be focused to achieve a sufficiently high luminosity. In a storage ring a bunch of accelerated particles is reused until the particle density has been reduced by many collisions, but a linac's colliding beams are 'thrown away' once they have passed through the interaction area. If particles with a given luminosity in a storage ring lose one per cent of their energy per circuit, a linear collider of equivalent energy efficiency must have a beam with 100 times the luminosity.

When tuning is completed in a few weeks, the positron and electron beams at the collision point will be in bunches 1.5 mm and 3 microns in diameter, each containing 70,000 million particles. With a repetition rate of 60 cycles per second, the scientists at SLC hope to create about 10 Z⁰ particles per day during the first year of operation. Later they want to increase the daily production rate to 300 Z⁰ particles.

Now that SLC construction is completed, the team at Stanford, under Nobel laureate Burton Richter, has begun to think about the next generation of linear colliders. They envision a pair of linacs, each about 5.5 km long, pointing at each other. They would be able to produce electrons and positrons at collision energies between 300 and 2,000 GeV.

Peggy Hellweg

UK industrial support slows

London

DESPITE repeated appeals from the government that British universities should look increasingly to the private sector for funds, the latest figures show that the rate of increase of income from industry and commerce halved last year. Provisional figures released last week by the University Grants Committee show that although university income from industry and commerce increased in 1985-86 by more than £11 million (or 24 per cent) to £59 million, the rate of increase is less than the 46 per cent scored in the previous year. As a proportion of the total income for specific research projects, the contribution from industry increased by only 0.5 per cent.

Many universities have set up special arrangements for attracting industrial in-

vestment. But the funds the sponsors offer frequently fail to account sufficiently for the indirect costs of research. "Many vice-chancellors believe that universities cannot afford to accept too many contract grants from industry because of the overheads", the committee spokesman said.

Last week's provisional figures show increases in other sources of non-government funds, with a 26 per cent increase in the contributions of private charities to £72.2 million.

Overseas sponsors, including companies and foreign governments, spent £28.5 million, an increase of a third on the previous year. Earnings from short courses, many of them specifically for industry, increased by a half to £39 million.

Simon Hadlington

British cancer fund to double its expenditure on research

London

WITH a 25 per cent increase in income last year, the UK Imperial Cancer Research Fund (ICRF) is in an expansionary mood. It plans to spend more than £500 million in the next ten years, by which time annual expenditure should have more than doubled. The fund is "bucking the trend" of UK research, says its director of research, Sir Walter Bodmer.

One almost immediate plan, for which a quarter of the fund's £32 million income last year was set aside, is to convert more of the administrative offices in the Lincoln's Inn Fields building in central London into laboratories. Protein structure and pharmacology research are likely to be the beneficiaries. But overall the balance of spending in the next decade will shift more in the direction of applied research as opportunities to move from the laboratory into clinical research are exploited.

While researchers are voting with their feet in response to the declining standard of facilities and uncompetitive salaries in government-financed biomedical institutes and university departments, ICRF

has no such problem, says Bodmer. ICRF cannot, of course, compete with the six-figure salaries with which US biotechnology companies lure the occasional scientist (Robert Kamen of Genetics Institute was an ICRF casualty of that type) but can afford the kind of salary and facilities to retain senior staff, claims Bodmer. This year's departure of three such staff (to start a Ludwig Institute, to a chair at the University of Oxford and to direct the Beatson Institute) is for the 'right' reasons, he says.

The ability of ICRF to maintain its protected position, let alone to expand as planned, depends on its success in persuading the public to donate money to cancer research. Legacy income is the most important single source and rose sharply last year, contributing £19.2 million towards the total of £32.5 million. But hopes of a sustained increase in income are largely pinned on a growing chain of shops around the country. These are predicted to bring in £14 million ten years hence. There is also some thought of reducing the £50 million or so accumulated capital fund that ICRF holds.

Peter Newmark

Radioactive release detected in Germany from Soviet source?

Munich

HIGHER levels of atmospheric radioactivity measured across Europe in the first half of March were most likely the result of a leak from a nuclear reactor in the Soviet Union. This suspicion, reported last week by the West German Institute for Atmospheric Research (IAR), has led the West German government to ask the Soviets for an explanation.

The Soviet Union has filed no report of an accident with the International Atomic Energy Agency (IAEA) in Vienna. But as the amount of radioactivity measured was so small, it would not be required to report it, said an IAEA spokesman.

Traces of radioactive iodine-131 and xenon-133 were first recorded in Finland and Sweden between 2 and 9 March. At the time, they were attributed to a nuclear test reportedly carried out by the Soviet Union on 26 February. But researchers at the IAR, where the increase in radioactivity was observed between 9 and 14 March, consider this implausible as xenon-133 has a half-life of just nine hours. The ratio of isotopes also rules out accidental release from a medical institution.

Based on the radioactivity levels measured in Germany (37 microbecquerel per m³ of air), the release may have amounted to only 1–10 curies, an amount "within normal permissible limits" for a nuclear reactor, according to an IAEA spokesman. (The accident at Chernobyl released 20 megacuries.)

One explanation for the release may have been a cracked fuel-rod canister which released these isotopes "up the stack" through the cooling system, hardly an uncommon occurrence, said the IAEA spokesman.

Steven Dickman

Vera Rich adds: The Soviet Foreign Press spokesman, Gennadii Gerasimov, the minister of nuclear power, Nikolai Lukonin, and the assistant chairman of the state committee for hydrometeorology and environmental protection, Valentin Sokolovskii, have all denied that there has been any leak of radiation from Soviet sources. The Soviet radioactivity monitoring network is working well, Gerasimov said at one of his regular press briefings on 14 April, and if there has been some sort of increase in radioactivity levels in the West, it is not of Soviet origin. □

Conservation treatment for breast cancer

London

THE first controlled trial of a new breast-conserving treatment for breast cancer has shown that it is as effective as mastectomy. On the basis of the trial, sponsored by the Imperial Cancer Research Fund at Guy's Hospital, London, the treatment will now be that of choice in Britain, according to Dr Ian Fentiman, deputy-director of the fund.

The therapy entails the removal of the tissues of the lump in the breast and of the lymph glands under the corresponding armpit. Radioactive iridium needles are inserted for 48 hours, delivering a dose of 20 gray, with external radiotherapy as a backup.

The trial involved 800 women, half of them mastectomized controls, over a period of five years. The subjects included 200 women treated at Guy's Hospital and an equal number in Belgium and the Netherlands. The results of the study have not yet been published, but it is reckoned to be significant because the rate of cancer recurrence is found to be the same in the subject and the control groups—roughly one per cent a year.

Fentiman says that the new treatment is applicable to patients with tumours up to 4-cm diameter, or to about 80 per cent of women with breast cancer in Britain, and that the cosmetic result has been excellent.

The death-rate from breast cancer in Britain is less than that in the United States, where pressure for conservation as an alternative to mastectomy has been mounting. Dr Samuel Hellman, physician-in-chief at Memorial Sloan Kettering Cancer Center in New York, the pioneer of the use of iridium implants, says that the new study should help to convince people that conservation techniques can be as effective as others.

Kathy Johnston

Italy to map the human genome?

Milan

Italy is planning to enter the race to map the human genome, chiefly because the National Research Council (CNR) believes it has a lot to offer. CNR president Professor Luigi Rossi Bernardini, announcing the decision, said that the council's mind had been made up as the result of a large meeting held in Rome under the chairmanship of Renato Dulbecco of the Salk Institute. The plan is to set up a new institute in Rome.

Paola de Paoli

Japan disputes need for moratorium on whaling

Tokyo

"WHALE resources are increasing. Why give up?" With these words emblazoned on her side, the mother ship of Japan's whaling fleet, the *Dai San Nisshin Maru*, sailed into Tokyo port last week, marking an end to Japan's commercial whaling in the Antarctic Ocean. But if the Fishery Agency has its way, the *Nisshin Maru* will soon set sail again for southern waters to carry out 'scientific whaling' for research purposes.

The International Whaling Commission (IWC) in 1982 called for a moratorium on all commercial whaling for five years from 1986 to 1990. Japan tried to side-step the ban by filing an objection to the commission's ruling but was confronted with a US threat to slash Japan's fishing quota in the US 200-mile exclusive economic zone. And in July last year, Japan withdrew its objection on the understanding that Japan's commercial whaling in Antarctic waters would end in 1987 and in Japanese coastal waters in March 1988.

The imposition of the moratorium has

raised bitter feelings in Japan, and over the past months the Japanese media have given a whale-by-whale account of the *Nisshin Maru*'s last commercial voyage. On 27 November, the crewmen cheered and broke open a cask of sake (rice wine) as the first minke whale was hauled aboard. The announcement on 14 March that the 1,941st minke whale, Japan's quota limit, had been caught was greeted by the crew with a more mixed response; those engaged in cutting up the catch said with deep emotion that they might never take up the cutting knives again.

Resentment in Japan over the moratorium runs high, not because whale meat forms an important part of the diet—today it is a delicacy eaten only on rare occasions—but because many Japanese feel that a traditional industry already on its last legs is being bludgeoned to death by Western environmentalists on emotional rather than scientific grounds. IWC estimates place the population of minke whales in the Antarctic Ocean at more than 300,000; whaling experts from the Fishery Agency say this population

is ample to sustain growth even if Japan continues whaling at present levels.

The issue seems set to come to the boil again in June at the IWC's next annual meeting, when Japan will put forward a plan to catch 825 minke whales and 50 sperm whales in the Antarctic Ocean in 1987–88 for scientific research purposes. The research whaling, to be carried out by the *Nisshin Maru* and two catcher vessels, will be supported by a budget of ¥350 million from the Ministry of Agriculture, Forestry and Fisheries.

A spokesman for the World Wildlife Fund has already condemned the Japanese move as a 'scam' because the proposed quota is close to half of Japan's 1986–87 commercial quota and the whale meat will be used for consumption after research.

Clearly Japan is hoping that its whaling industry will be able to limp on and amass enough scientific data to convince the IWC that the moratorium should be lifted when it comes up for review in 1990. But one Tokyo restaurant specializing in whale meat has set aside storage space in a freezer for 70 tonnes of whale meat—the equivalent of about 14 minke whales—enough to keep the restaurant in business for at least the next six years.

David Swinbanks

Superconducting ceramics (contd)

Tokyo

HARDLY a day goes by without a Japanese company announcing a 'breakthrough' in the industrial development of the new superconducting ceramics. Now it is the turn of NEC Corporation.

It is less than a month since NEC researchers reported at Nagoya that they had observed the Josephson effect in the new ceramics at 80 K (see Nature 326,432; 1987). But already they claim to have a functioning Josephson junction device. Measuring 1 by 1 by 10 mm, the junction is composed of two layers of single-phase yttrium-barium-copper oxide. Both AC and DC Josephson effects have been observed in the device at temperatures ranging from 0 to 90 K.

Ultimately, NEC hopes that such devices will be used to make a superfast computer that can operate at liquid nitrogen temperatures. But temperature is not the only problem besetting the development of Josephson junctions into computers. Nobody has succeeded in combining several such devices into a viable chip because of different instability effects in each junction.

Details of NEC's new device will be announced at the Material Research Society spring meeting to be held in the United States from 21 April. David Swinbanks

Explicit AIDS campaign offends TV watchers in Australia

Sydney

IN Australia, an AIDS (acquired immune deficiency syndrome) campaign with a touch of mediaeval horror to it is causing controversy. The centrepiece of the A\$3 million national campaign is a television advertisement featuring the Grim Reaper as the symbol of AIDS death. The advertisement is designed to shock and scare. A rotting corpse clad in rags, carrying a scythe, aims a bowling ball at a group of ordinary Australians set up as tenpins. They are knocked dead. A little girl clutching a doll and crying is knocked flying for a spare.

The message of the campaign is that it is not only intravenous drug-users and homosexuals who are at risk from AIDS and that the only way of avoiding such a fate is to engage in 'safe sex'.

There have been complaints that the advertisement gives children nightmares, and demands that it not be screened when they could be watching.

Criticisms of the campaign include the lack of clear information on how AIDS is actually transmitted, that it ignores the importance of a partner's sexual history over the previous five years, and the use of exaggerated figures such as the estimate that "there are already 50,000 suspected carriers of the AIDS virus in

Australia"; other estimates put the number at more like 3,000. So far, there have been 238 deaths from AIDS in Australia.

According to the Minister for Health, Dr Neal Blewett, the shock strategy of the advertisement was designed to capture people's attention in competition with the vast amount of information pouring out of the media daily, to emphasize the seriousness of the problem and to encourage people to make further inquiries for themselves. In this, it seems to have been successful. After the launch of the campaign, the number of calls received by Sydney's main AIDS clinic quadrupled.

The AIDS campaign is also causing problems for schools. In response to the AIDS campaign, one of the national current affairs programmes intended for schools prepared an edition explaining what it all meant. But the New South Wales Department of Education, which is still deciding how to handle the question of AIDS, decided on the day before the screening that students would need parental permission to be allowed to watch it. Because of the lack of time to inform parents and receive a reply, very few children in the state saw the programme.

Charles Morgan

India launches new plan for unconventional energy sources

New Delhi

AN ambitious plan to produce by the end of the century 15,000 MW of electricity in India from sunlight, biomass, wind and microhydroelectric plants has been drawn up by the Department of Non-conventional Energy Sources (DNES). The target is half as much again as that set for nuclear electricity, but the plan has yet to win government approval.

One feature of the new plan is that decentralized energy systems would be established throughout India to meet local demands for fuelwood, cooking gas and irrigation power. According to DNES secretary Maheshwar Dayal, total energy generation from non-conventional sources by the year 2001 would be equivalent to 250 million tonnes of coal, or one-fifth of the total energy needs of India. Dayal says that the scheme "is well within the realm of practical possibility". The cost of the scheme is estimated at \$40,000 million, of which the government would put up 24 per cent and the rest would be palmed off on industry and energy cooperatives.

Under the DNES plan, about 2.5 million hectares of land will be brought under energy plantations coupled with power-generating units based on gasifiers and sterling engines. Energy crops on 1,000 hectares can generate 3 MW of power, and a total of 6,000 MW would be produced by the biomass project. Part of

the harvest would be for fuelwood.

DNES has estimated that more than 20,000 MW of potential windpower exists in the coastal areas. It is proposed to generate 5,000 MW of power from large-scale wind farms and another 2,000 MW from solar thermal power stations. Some 1,000 micro-hydel stations will provide an additional 2,000 MW.

The plan also provides for 12 million big biogas plants to provide cooking gas, 50,000 solar photovoltaic and 50,000 wind pumps for lift irrigation and drinking water supply. Sewage from all major cities and towns will be treated and 50 MW of power produced from the sludge. Energy will be recovered from the effluent of all the 140 distilleries in the country and 40 projects are to be set up for recovering 160 MW of power from solid municipal wastes. To save fuelwood, 100 million improved wood-burning stoves would be installed in rural homes to

replace the inefficient 'chulahs' now in use.

Within the next three years, 5,000 villages are to be made self-sufficient in energy, using a combination of energy systems and devices. A Renewable Energy Development Agency has already been set up to finance, through soft loans, a large number of the non-conventional energy projects.

Photovoltaic modules are now manufactured in India and the raw material, pure silicon, is also produced indigenously on a commercial scale. But a good number of the projects under the DNES plan are likely to open a vast market for vendors from abroad. Present Indian projects for a solar thermal power station are based on US and West German technology. The five wind farms in operation use models developed by Denmark and the Netherlands. The small turbines needed for micro-hydel stations are not made in the country and India's first refuse incineration plant in New Delhi, which will generate 3.75 MW of power from solid waste, is being set up by Denmark.

K.S. Jayaraman

Chinese intellectuals need not fear fall from Party favour

London

FEARS that the 'reorganization' of the Chinese University of Science and Technology last January (see *Nature* 325, 290; 1987) signalled the return of hard times for China's intellectuals seem to have been unfounded. The removal of university president Guan Weiyan, and vice-president Fang Lizhi, and the expulsion of the latter from the party, has not meant the termination of their scientific careers. Fang's "offence" was a serious one; he had allegedly advocated "bourgeois liberalism" and "total Westernization". A few days after his dismissal, the party general secretary, Hu Yaobang, who admitted favouring Fang's views in a self-criticism, had to resign.

Fang and Guan were immediately found posts by the Chinese Academy of Sciences, Guan in its Institute of Physics and Fang at the Beijing Observatory. Soon after, when the president of the Chinese Academy of Sciences, Lu Jiaxi, retired, the new president, Zhou Guangzhao, made it clear that the current drive against "bourgeois liberalization" did not mean absolute rigidity of outlook among academy personnel. "The academy will . . . allow different views. Debates can lead to truth", he said. And, he added "we still allow scientists like Fang Lizhi to work in the academy and present academic reports. The academy has not limited Fang's per-

sonal or scientific activities."

One exponent of a more relaxed attitude to diversity of views among academics seems to be the prime minister, Zhao Ziyang, who since the resignation of Hu Yaobang has also been acting general secretary of the party. Addressing a group of delegates from the academy recently, he said that opposition to "bourgeois liberalization" should not develop at the expense of democratization, and that research institutes, in particular, should be one step ahead in democratization. Channels for discussion and dialogue between leading cadres and intellectuals must be opened up, he said, so that the intellectuals can have outlets for their views.

So long as Fang and Guan stick to research, it seems, the leadership is prepared to overlook recent political errors. But at the time of his dismissal, Fang was scheduled to visit Britain, to attend the Newton Conference in Cambridge in July. The conference organizers have had no communication from him since his acceptance of their invitation, which was dated in late December, before his dismissal. Recently, however, the Hong Kong journal *Pai Hsing* reported that Fang's trip to Britain will take place as scheduled, and that this is the personal decision of Zhao Ziyang. If Fang does indeed arrive in Cambridge, it will confirm Zhao Ziyang's assurances.

Vera Rich

Nuclear plants spark protest

Bangalore

ENVIRONMENTALISTS in the southern Indian state of Karnataka are up in arms at the central government's plans to build four nuclear power reactors at the coastal site of Kaiga. The intention is that two plants of 235-MW capacity should soon be built there, to be followed by two 500-MW plants at a later stage. The reactors will be based on the Canadian CANDU design, with natural uranium fuel and heavy water as a coolant.

One complaint is that the dense tropical forests surrounding Kaiga are an important natural resource, but the region is also seismically active. Other objectors point out that the proposed nuclear site is only 30 km away from India's largest naval base at Karwar, which could make it vulnerable to enemy attacks. Although the people of the region are mostly illiterate, environmental groups have made sure that they now understand the potential dangers of nuclear plants.

Radakrishna Rao

Striking the balance on AIDS

SIR—The “Sombre view of AIDS” proposed by Malcolm Rees (*Nature* 326, 343–345; 1987) is based on the suggestion of a 15-year mean incubation for the human immunodeficiency virus (HIV) transferred by blood transfusion and a consequentially high estimate of future incidence. He uses Peterman’s data (given in Rees’s Table 1). His mode of analysis is perhaps too indirect. At all events the simple ‘actuarial’ approach leads to a different conclusion.

Write p_s for the chance of AIDS (acquired immune deficiency syndrome) diagnosis in the s -th year after infection. The ratio $r_s = p_{s+1}/p_s$ may be directly estimated from the diagonal totals in Table 1 relating to the same set of annual cohorts of infecteds. The estimates (for the five values $s=1\ldots 5$ whose measurement the data permits) are:

5.00, 1.77, 1.59, 0.84, 1.18

This is not consonant with a gaussian distribution with $\mu=15, \sigma=5$ which gives values:

1.73, 1.67, 1.63, 1.56, 1.47

The observed ratios suggest the probabilities rise rapidly to a peak around five years. Rees’s suggestion of a log-normal graduation could give rise to ratios with this characteristic. So equally would graduation using other distributions like the gamma (which arises for example as a special case of the analogous distributions advocated by Julian Peto in carcinogenesis using staged models). Whatever form is used for graduation the salient feature is the same: a mean incubation of about five years rather than fifteen. This leads to an order of magnitude reduction in Rees’s “sombre” projections.

D.E. BARTON

Department of Computer
Science and Statistics,
Queen Mary College,
University of London,
Mile End Road,
London E1 4NS, UK

SIR—Your report (*Nature* 324, 611; 1986) on AIDS in Africa highlights “the sorry state of affairs that exists between French and US AIDS researchers” and a possible “African origin of AIDS”. It seems to us that such perceived problems are unduly emphasized by an overzealous press, further complicating attempts to conduct important research. We have, for example, maintained an international agreement for a trilateral study on AIDS-related viruses with both French and US investigators. This arrangement has been focused on common goals, based on mutual respect, and has been free of problems related to nationalism.

Perhaps a more serious concern for Africa is the need for careful and objective

evaluations of experimental results that could have major implications for public health policies. Some of the first sero-epidemiological studies conducted in Africa reported large proportions of ELISA (enzyme-linked immunosorbent assay) test results that were found to be false positives. This experience may have helped to create a climate of reluctance to accept the sub-sequent results that accurately revealed the extent of the epidemic in Central and East Africa.

In West Africa we now have a situation where a significant proportion of people have been exposed to a different class of retrovirus that is highly related to the primate retrovirus, STLV-3, and variously designated HTLV-IV, LAV-2 and HIV-2. Premature conclusions that this new class of retrovirus is as pathogenic as HIV-1 could result in unsound policy decisions in various West African countries. It is particularly important that appropriately controlled population-based studies be conducted to address the possible link between infection with this new class of retrovirus and AIDS or any other pathology. Such studies should be conducted as soon as possible and should incorporate highly specific assays that allow one to distinguish between HIV-1 and HIV-2. Only then can a sensible and sound health policy response be formulated.

S. MBOUP, D. RICARD, C. BOYE,
A. GAYE, J.L. SANKALE

Laboratoire de Bactériologie-Virologie,
Faculté de Médecine et de
Pharmacie de Dakar,
Dakar-Fann, Senegal

I. NDOYE

Bureau des MST,
Ministère de la Santé,
Dakar, Senegal

L. SANGARE

BP 8175, Ouadougou, Burkina-Faso

F. MOWONDI

BP 08, Foumban, Cameroun

B. DIALLO

BP 217 bis, Conakry, Guinée

M. RAJAONARIVOY

BP 460 Antananarivo, Madagascar

Singapore maligned

SIR—A letter under the headline “Dissent disallowed in Singapore” (*Nature* 324, 298; 1986) contained factual inaccuracies and appears to be an unprovoked attempt to discredit Singapore before the international scientific community. The Fourth Federation of Asian and Oceanic Biochemists Congress that took place in Singapore was the most recent in a series of triennial regional meetings organized by biochemists of the Asia-Pacific region. As such, its planning and eventual success as a scientific meeting did not depend on

aggressively seeking the participation of the international scientific community. That this meeting attracted about 700 registrants from 33 countries is perhaps the most telling evidence that scientists are not facile targets for alarmists.

The congress was organized independently of the planning of the Institute of Molecular and Cell Biology in Singapore. No attempt was made to recruit scientific staff for this institute at the meeting.

The statement that “the voice of dissent in Singapore... is to be disallowed at all cost” is false and invidious as is the grotesque suggestion that scientists elsewhere should use their influence to rescue the people of Singapore from their unpleasant fate.

EDWARD JACOB

163, Sixth Avenue,
Singapore 1027

Papal IVF poser

SIR—Your leading article on the Vatican’s pronouncements on the ethics of *in vitro* fertilization (*Nature* 326, 229–230; 1987) leads me to wonder if the unfortunate couples who have been determined to be mutually infertile have the right in God’s eyes to engage in intercourse? After all, it could be considered that ‘science’ has determined that there is no procreative future in such an act, and so the kind permission to engage in “natural acts which are intrinsically directed towards procreation” must most regrettably be withheld. If only scientists were unable to diagnose sterility perhaps these couples would at least have the right to keep on trying.

Nostra maxima culpa!

T.B. GREENLAND

39, rue St Jean,
69005 Lyon, France

Three degrees

SIR—In K.S. Jayaraman’s article on “Superconductivity” (*Nature* 326, 237; 1987), there are some factual errors.

(1) There are several teams working on superconductivity in India and the team that reported, following the work of Chu *et al.*, the highest superconducting onset and zero-resistance temperature for the first time is solely from the Tata Institute of Fundamental Research (TIFR). No scientist from the Bhabha Atomic Research Centre (BARC) is involved.

(2) The zero resistance state is reached at 87 K and not at 90 K.

(3) This is not the same team as the one that a month ago reported superconductivity in a lanthanum oxide system (which is of course a BARC–TIFR team).

E.V. SAMPATHKUMARAN

Solid State Physics Group,
Tata Institute of Fundamental Research,
Bombay 400 005, India

Access to drugs on clinical trial

SIR—I was amused by Dr Vincent T. DeVita's comments reported in your article on Biotherapeutics Inc. (*Nature* 325, 99; 1987). Many patients who have attempted to take part in National Cancer Institute (NCI) trials would dispute that these new treatments are being made "available to the public as fast as is prudent". Clearly, there are many new approaches available for the treatment of cancer and these are *not* being made quickly available. There are many patients who wish to try experimental treatment who are unable to gain access to such approaches because of the restrictive nature of the regulated monopoly administered by the government/academic sector. *Nature* should poll patients with respect to DeVita's claim.

It is puzzling to continue to see the word ethics coming from the NCI. Patients do fund research at Biotherapeutics but the public should know that patients also pay for experimental approaches in the NCI/university system. When I was director of the Biological Response Modifiers Program for DeVita, we billed insurance companies for the cost of hospital care at Frederick Memorial Hospital while patients received experimental treatment with interferon and other new biological approaches. Thus, the patients, through their insurance company, clearly "paid for experimental treatment" in that NCI study. In addition, it is well known that university researchers allow and encourage their hospitals to bill insurance companies and sometimes patients for the cost of hospitalization when patients are on NCI/university experimental protocols. Many of these trials are initiated and monitored by the NCI and experimental drugs made available for these trials through the NCI. Thus, patients and their insurance companies are paying for the clinical costs in these experimental trials. According to DeVita's standard, is not that unethical?

ROBERT K. OLDHAM

Biological Therapy Institute,
Riverside Drive,
Franklin, Tennessee 37064, USA

Problems facing university science

SIR—The recent pay awards agreed by the Association of University Teachers favour those at the top end of the scientific profession and will not help in attracting the younger researchers on whom the future of science in Britain ultimately rests. To do a PhD to enter university research in Britain these days is to condemn oneself to years of penury and constant insecurity. When senior scientists next request more

funding for basic research (and basic jobs), the government's response will be that money was made available and that it went straight into the pockets of the applicants, already overpaid in the prime minister's eyes, assuming that her recent statements about cutting top rates of income tax to stem the brain drain are to be taken at face value. The voting public will offer science at the universities little sympathy, even if we bend over backwards to show how useful and applied, cheap and streamlined we have become.

We will not prevent the erosion of British university science by permitting the award of such selective pay increases and by aiding and abetting the government in its plans for rationalization of research and teaching. It is little use pretending that we can be 'great value for money' in the short term. We cannot be trained to turn basic research into pounds sterling within the life of a government or two. This particular government, concerned as it is with value for money and the quick turnover of the marketplace, knows that basic research is expensive. It also knows that financial returns on the investment can be enormous but often take decades to appear. The pursuit of basic research, the task for which the universities are best equipped, is thus anathema to Thatcherism and the present government is relatively unconcerned when we leave to ply our expensive trade abroad. Britain's loss is deferred to the 'never-never'. We must remind the electorate and political parties of the necessity of investment in university research, despite its long-term objectives, refuse to assist in the removal of the basic research and teaching roles of the universities and, most importantly, ensure that the little money cast in our direction is not spent on pay rises which merely allow those remaining in British science to watch impotently over its demise in slightly greater comfort.

SIMON P. WOLFF

Department of Clinical Pharmacology,
School of Medicine,
University College London,
5 University Street,
London WC1E 6JJ, UK

SIR—The recently published Report to the House of Lords of the Select Committee on Science and Technology entitled *Civil Research and Development* (HL 20-1, HMSO) recommends that the "University Grants Committee selectivity exercise should be repeated in less than five years and the process be more open in future".

Not only is the openness important, but also the guarantee that like will be compared with like. This department, which was returned as Cost Centre 8 (other studies allied to medicine) was rated as average. It has research grants support of £1,817,332; has published 195 papers in the past four years; has pioneered medical

imaging; has staff who are Royal Society Wellcome Foundation Gold Medallists and prize-winners in the BTG Academic Enterprise Competition; has spawned two companies; has one of the only two cyclotrons in the United Kingdom for medical imaging research; and has 15 PhD students and 28 MSc students. All this with an academic staff of five whole-time equivalent university posts and 40 graduate staff including National Health Service (NHS) and research grant staff.

I endeavoured to discover those departments that are above average and outstanding, so that we could learn how to improve our performance. According to *The Times* of 3 June 1986, the others graded in Cost Centre 8 (the London medical schools were excluded for reasons of space) were:

Outstanding: Aston; UMIST. Above average: King's. Average: Keele; Aberdeen; UWIST. Below average: Bradford; City; Exeter; Hull; Liverpool.

A short session perusing the *Commonwealth Universities Year Book 1985* showed me that the only departments in these universities that might reasonably have been grouped into Cost Centre 8 were named as follows: ophthalmic optics (3); optometry (2); postgraduate medical (2); environmental studies (2); nursing studies (2); communications and neurosciences (2); clinical communication (1); community health (1); health studies (1); physical education (1); medical sciences (1); biomedical engineering (1); medical physics (1).

If the UGC selectivity exercise really did compare these departments, then its absurdity is obvious, even to those unfamiliar with our interdisciplinary field which usually has active service commitments to the NHS as well. This is reinforced by the fact that those universities known to us personally as active in our field are not mentioned in *The Times* list, for example Surrey, Leeds, Sheffield.

I cannot accept the verdict of average for this department unless I can see openly that it has been done fairly. It is my belief that it may already have led to the loss of a medical-imaging research grant of £470,400 to the department.

The Select Committee reports, in addition, that it "...shares Sir Peter Swinnerton-Dyer's reluctance to establish within the UGC an elaborate assessment bureaucracy...". However, surely *someone* must be able to ensure that like is compared with like, and enable it to be seen to be operating fairly? Does it have to be elaborate?

JOHN MALLARD

Department of Bio-Medical
Physics & Bio-Engineering,
University of Aberdeen and
Grampian Health Board,
Foresterhill,
Aberdeen AB9 2ZD, UK

Fraud versus carelessness

SIR—Stewart and Feder¹ protest a little too much, thereby detracting from the valuable comments they make.

Mistakes have often been made and will continue to be made, even by the eminent. To quote Heilbron²: "Sometimes, to be sure, Rutherford's semi-quantitative results misled, as in the conclusion, supported by a 20 per cent error, that positive and negative ions in a gas drift at about the same velocity". They only detract from the credibility of the author and, if repeated too often, will simply "be ignored rather than deliberately gainsaid"³. As scientists are held to engage in a quest for credibility⁴, this may itself be penalty enough. Do mistakes add up to fraud? How easy is it anyway to fabricate data?

I recently participated in a graduate course in pharmacology in which students were required to make several oral presentations. I asked them to imagine that they were promoting a novel compound for the management of rheumatoid arthritis in a country that had suddenly opened its doors to Western medical technology, taking as their starting point existing information on anti-inflammatory drugs, gold compounds, immunosuppressants and so on. They were told their compound should purport to be an improvement on existing drugs. The students performed brilliantly. They concocted variations on existing chemical structures, gave attractive names to their compounds, provided pharmacological and toxicological profiles, devised double-blind randomized clinical trials with appropriate criteria for selection of patients and assessment procedures and cleverly interpolated their wholly fictitious data into existing information. As an observer, I was impressed with their performance and utterly shocked by how real it all seemed. They could well be reviewing published information or could have even sent their data in. It all seemed so facile. The point to ponder is not that fraud occurs, but that it appears to be so uncommon — at least till now.

But will the frequency of fraudulent publications increase? This is the rub. Will carelessness, repetitive publications and gamesmanship lead cynically to fraud? Do these form a continuum? It is particularly disturbing that in recent years the highly publicized accounts of fraudulent work have been linked to scientists in the biomedical disciplines, particularly in North America.

The explosive growth of the biomedical community has come at a price. Universities have allowed unbridled growth of their medical schools, as medical faculties attract funds, often generate their own salaries and plough money back into the

universities. The young clinician-scientist is forced to juggle his time between seeing patients, generating his salary, administering his unit, teaching students and doing research. Because of ceilings on incomes, clinicians attached to medical schools often earn far less than their peers in practice and thus make considerable financial sacrifices to pursue their careers. The powers-that-be who indulge in gratuitous quantification force them to publish more than they should.

In all this, senior clinician-scientists are squarely to blame. They cannot take refuge in describing fraud as aberrant behaviour, as Braunwald has done, for they create the environment that makes fraud possible. Macklem⁵ noted that "the cut-throat and poisoned atmosphere that results is hardly conducive to originality and creativity". They exercise considerable influence on granting agencies, sit on tenure and promotion committees and often become role-models to be emulated. In fact, many are; I have personally been privileged to work under several. A simple message that excessive publications will be penalized will be easily understood. Senior scientists (particularly in clinical disciplines) can do far better for their junior colleagues by protecting their research time and cautioning them that incessant brownian motion does not necessarily equate with progress. Frankly, it is not too difficult to be productive in a quantitative sense — it merely requires a certain lack of imagination and a willingness to get bored and be boring. About 20 years ago, Ziman wrote of those who "rush into print with the notebooks of the previous day's experiment"³ (intellectual fraud?). It is only one small step then to rush into print with non-existent notebooks of never-done experiments.

P.K. RANGACHARI

*Department of Medicine,
Intestinal Disease Research Unit,
McMaster University,
1200 Main Street West,
Hamilton,
Ontario L8N 3ZS, Canada*

1. Stewart, W.W. & Feder, N. *Nature* **325**, 207–214 (1987).
2. Heilbron, J.L. in *Rutherford and Physics at the Turn of the Century* (eds Bunge, M. & Shea, W.R.) (1979).
3. Ziman, J.M. *Public Knowledge*, p.115 (Cambridge University Press, 1968).
4. Latour, B., Woolgar, B. in *Science in Context* (ed. Barnes, B. & Edge, D.), 35–43 (Open University Press, Milton Keynes, 1982).
5. Macklem, P.T. *Clin. Invest. Med.* **8**, 85–88 (1985).

SIR—In your issue of 15 January 1987, your leading article and Commentaries by W. W. Stewart and N. Feder and E. Braunwald deal with a scientific fraud, which the whole community would agree was reprehensible. However, other practices occur much more widely, but they are rarely admitted. Among these, some scientists:

(1) do not attempt to publish results of their own or their team's experiments that

contradict their own hypotheses;
(2) do not carry out crucial control experiments, which they have no reason to believe have already been published, or to which their attention has been drawn;
(3) pay lip service to the known limitations, artefacts and assumptions inherent in the techniques they have employed, but do not take them into account when drawing conclusions or making calculations in respect of their experiments;
(4) accept or demand co-authorship of publications — thereby assuming equal intellectual responsibility for publications resulting from experiments — to which they have given little or no supervision and in which they have hardly participated physically;
(5) omit from publications important details of experiments that would be necessary to make them repeatable;
(6) do not cite publications reporting experiments or observations that predate or contradict their own;
(7) as referees, give insufficient attention to papers in relation to the time and resources the authors have employed, and then hide behind their anonymity when criticized;
(8) as referees, recommend the rejection for publication of manuscripts presenting views or findings that they do not like;
(9) as referees, recommend the rejection for publication of manuscripts whose conclusions they *feel* are wrong, but in which they cannot identify the mistakes;
(10) as organizers of conferences, seminars and symposia, prevent the presentation of views, however well argued, that they consider 'heretical', 'controversial' or 'unorthodox';
(11) are unwilling to discuss their own research work or fundamental aspects of their own work or discipline, privately, publicly or by correspondence;
(12) influence committees to reject applications for grants likely to produce results they would not like, or by research workers of whom they disapprove.

I believe that the scientific community is far too tolerant of such practices, often dismissing evidence for them by pointing out that research workers are as human as the rest of the population. I believe that the combined impact of the above practices upon the corpus of knowledge is far greater than the rare and unpopular acts of fraud.

H. HILLMAN

*Unity Laboratory of Applied
Neurobiology,
University of Surrey,
Guildford,
Surrey GU2 5XH, UK*

Letters submitted for Correspondence should be typed, double-spaced, on one side of the paper only.

Should technology be scorned?

Science and technology are parts of a seamless spectrum, but there is a good reason to believe that the literature of engineering does a disservice to the professional community it purports to serve.

TECHNOLOGY is largely, but not wholly, the product of scientific enquiry, so why is the first so isolated from the second? The question is not as academic as it may seem. In Britain, for example, governments talk as if the gulf is the cause of Britain's poor economic performance. Engineers also often have a sense of being poor cousins, sharpened by the familiar observation that successful projects (space launches, for example) are called "scientific" triumphs, but failures, "engineering" disasters.

Historically (and recently), the formation of societies such as the National Academy of Engineering in the United States and the Fellowship of Engineering in Britain is partly a response to grumbling of this kind. It may be a more serious, but misguided, matter that grant-making agencies seem bent on blessing university engineering departments with research funds on the pattern familiar in science departments when there is no knowing whether this is how academic engineering would be best enlivened, in teaching, discovery or by means of more direct contributions to economic prosperity.

Part of the difficulty is that engineers are their own worst enemies when, as some do, they seek to endow their profession with a mystique that is essentially divisive. The argument is that engineers differ from scientists in their ambition to get things done, not just to think about them. This, for example, was one of the themes of the Finistion report on British engineering education which, seven years ago, set the foundations for the pattern of university engineering now emerging. It may be more relevant that the mystique is externally imposed by the usual requirement that engineering courses at universities must be validated by professional organizations. (The same applies to medicine, the oldest form of applied biology.)

That many successful engineers (from the two Stephensons and Brunel to Remington and Edison) were impatient of abstraction goes without saying. It is also usually the case that practising engineers must deal with problems of a practical nature that do not often arise in research laboratories — labour relations, economics and the law, for example. (A sufficient recognition of that would make people as respectful as engineers could wish.) But some kinds of engineering are indistinguishable from applied mathematics, while many research laboratories contain pieces of home-built equipment

which, by their function, precision and complexity, would not shame a professional engineer. The mere sight of, say, a high-energy physics particle detector should make that plain.

What this implies is that engineering and the practice of science are a seamless spectrum. If technology is the hard substance of engineering (labour relations and the law excluded), there should be no recognizable boundary between that and research proper. But may not people's objectives, and thus the nature of the intellectual work they do, differ profoundly? Although both engineers and, say, physicists may build similar pieces of equipment, are not their motives very different? Superficially, one might say that scientists' equipment exists so as the better to understand natural phenomena while that built by engineers is directed at some useful task external to itself.

That argument would be the more convincing if it were not that engineers, having built a new machine, find themselves impelled towards better explanations of its functions. The machine becomes the phenomenon, turning engineers into scientists. Those who built the first steam-engines found it necessary quickly to construct steam tables, essentially empirical substitutes for a full thermodynamic understanding of the dependence of pressure on temperature and density (but any other pair of independent variables will suffice) for two-phase water. Modern engineers, software engineers for example, are impelled in similar directions by the need to generalize their experience of a first machine (which may be a program) to the design of its successors.

Naturally, people in this state of mind are not so foolish as to pretend that everything they need to know can be deduced from first principles. Like other scientists, biologists for example, engineers are used to working with empirical generalizations. Modestly, perhaps because they have a good sense of the high cost of empirical measurement, they rarely dignify their generalizations as "laws" in the sense of Newton. Sadly, but not necessarily out of a sense of modesty, they tend to hug to themselves whatever generalizations they may have formed in their own minds. One passing result is that the professional literature of engineering is a disgrace: an outside innocent cannot hope by any amount of diligent reading to know what is in its authors' minds.

The best-known of the civil engineering journals, for example, contain descriptions of the largest earthwork dams; lesser journals describe smaller structures. Other people wishing to build dams on rivers of arbitrary size will have to make their own interpolations between contrasting circumstances. Much the same is true of the rest of applied science, even of medicine: readers are supposed to be able to accumulate case-records in their heads until they arrive at their own understanding.

The pity of this state of affairs is that the reality is much more interesting. Engineers in the real world grapple with two problems, of which the chief might be called that of optimization; what is the best (cheapest, most energy-efficient, fastest) machine to do some job. Extrapolating from known machines can yield great benefits, but the big prizes go to those who sense when it is time to change the mould. (*Canard* (wings at the front) aircraft will one day make sense, but it will be a clever person who knows how and when.) Might it not help us all if these questions were taken up directly in the literature of engineering?

The other absorbing question is that of complexity, or of how one comes to terms with one's use of a machine for which there is no *ab initio* explanation. Real life is amply full of demonstrations that one can survive: many of those who do not understand why lead-free gasoline tends to knock have trouble-free driving licences. The fact that mathematicians have so far failed to provide a finite solution of the general travelling-salesman problem, that of how most economically to travel some path between an arbitrary number of destinations, may be a sign that the need to survive without full understanding is perpetual.

Under both these heads, not to mention that of the mutual prosperity, engineering bursts with interest. Why is the literature of engineering so arid? To fear that many people have been restrained from writing what they might have written by the knowledge that their patent applications have not yet been approved would be venal, but it is a constant disappointment that even the accounts of novel semiconductor devices tumbling from the pages of *Applied Physics Letters* are as laconic as appears to be the rule. Could it be that engineers are their own worst enemy in this more serious sense?

John Maddox

Antigen presentation

Fugue in T-lymphocyte recognition

Ronald H. Schwartz

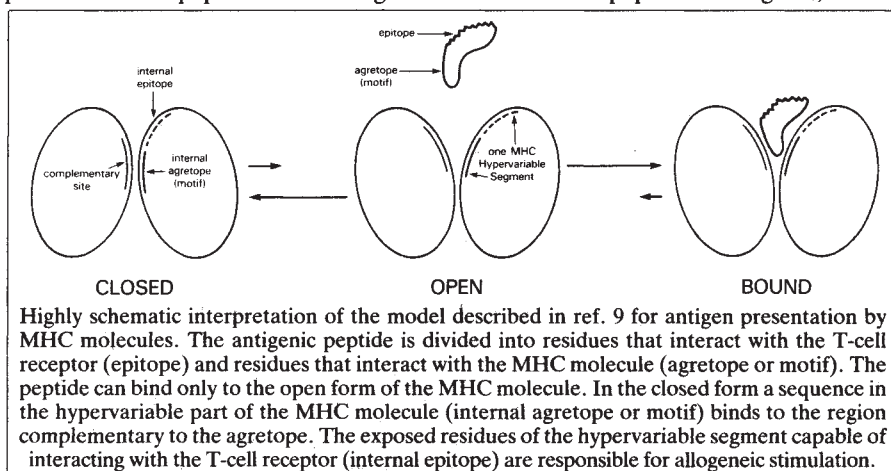
Two recent News and Views articles have described how the recognition of foreign antigens by T lymphocytes is mediated by a cell surface receptor¹ that recognizes two ligands, a peptide derived from the antigen and a glycoprotein called a histocompatibility (MHC) molecule located on the surface of a second interacting cell². Early functional experiments suggested that all three molecules, the T-cell receptor, the antigenic peptide and the MHC molecule, interact in a trimolecular complex to initiate T-lymphocyte activation³. Subsequent biochemical studies demonstrated binding of pairs of these reactants: the antigen with the MHC molecule⁴ and the T-cell receptor with the antigen⁵ (although the latter could be seen only if the antigen was studied in a polymeric form). Current excitement arises from further studies on the antigen-MHC molecule interaction. Buus *et al.*^{6,7} now show that the equilibrium constant is a consequence of very slow on- and off-rates, thus revealing an unexpected stability of the complexes. In addition experiments by Guillet *et al.*⁸ suggest that each MHC molecule has only a single combining site. Finally, analysis of the primary sequences of all peptides binding to the same MHC molecule reveals a related pattern of amino acids (motif)⁹. Surprisingly, this motif is also found in the MHC molecule binding the peptides, suggesting that the foreign peptide displaces an internal peptide sequence when it binds.

Foreign antigens recognized by T lymphocytes are almost always fragments of proteins. These peptides have been subdivided experimentally into two components, the amino-acid residues that interact with the T-cell receptor (epitope) and the amino-acid residues that interact with the MHC molecule (agretope)³. Agretopes do not influence the specificity of the stimulated T-cell clones, rather they determine the ability of an individual's immune system initially to recognize the peptide as foreign¹⁰. Functional studies first suggested that the agretope and epitope could behave independently³, and this has now been elegantly demonstrated by Guillet *et al.*⁸. These authors identified three distinct peptides from different proteins that interact with the same MHC molecule. Although T cells specific for one peptide could not be stimulated by the other two, these two peptides could competitively inhibit T-cell stimulation by the first peptide, suggesting that the three peptides were binding to identical or overlapping sites on the MHC molecule. To test this hypothesis, new peptides were made con-

taining only a few critical residues of the epitope of one peptide with the agretope of either of the other two. These 'hybrid' peptides are capable of stimulating T cells specific for the peptide donating the epitope residues.

The ability to scramble epitopes and agretopes led Guillet *et al.*⁹ to examine carefully the primary sequences of all sets of peptides interacting with the same MHC molecule. They found very similar amino-acid patterns in each set, which they called motifs. L/V, E, -, -, R/K, R/K, -, -, A/V, -, -, K, for example, is a motif present in three peptides interacting with

tion (discussed below), they found that the peptides show the greatest binding to that MHC molecule used by the animal to make an immune response. In addition, they proved that the binding they were measuring is biologically relevant by isolating the complex, inserting it into liposomes on glass beads and using this material to stimulate T lymphocytes. Their most intriguing observation was the discovery that the association rates for the reaction are about $1 \text{ M}^{-1} \text{ s}^{-1}$, about 10^3 -fold lower than typical antibody-antigen reaction rates in solution, which are often only diffusion- and orientation-limited. This result raises the possibility that very few MHC molecules are in the right conformation for binding the peptide. This could imply that the MHC molecules exist in two states (see figure), a closed conformation in which the internal motif is bound to the peptide binding site, and an



the MHC molecule E^d (single-letter code for amino-acid residues). These motifs could represent the structural basis for agretopes. The most surprising result was that almost all the elements of the motif are also found in a portion of the E^d molecule called the hypervariable region, the part of the molecule that differs from individual to individual. This observation is unlikely to be coincidental (although no rigorous statistical analysis has yet been done). Guillet *et al.* suggest that these residues in the MHC molecule bind to a complementary site in the same molecule. It is this site to which the foreign peptide binds, via its agretope, presumably by displacing the internal motif (see figure).

How can a free peptide displace an integral segment of the MHC molecule? The answer to this could lie in the recent physical measurements of Buus *et al.*^{6,7}. In following up the initial peptide-MHC molecule binding studies of Babbitt *et al.*⁴, they discovered that the peptide dissociation rate is extremely slow for a reaction with a dissociation constant of 1 to 10 μM . This allowed them to use gel-filtration methods to separate bound from free peptide and to assess the binding kinetics for 12 different peptides. With one excep-

open conformation, free to bind the peptide. Peptide binding would stabilize the open conformation.

In such a model, peptide binding to the MHC molecule would be the rate-limiting step in formation of the trimolecular complex because of the limited availability of MHC molecules in the open configuration. This, of course, assumes that a T-cell receptor is around to recognize the antigen-MHC molecule complex once it forms. Interestingly, Guillet *et al.*⁹ found one anomalous result in which a peptide from the λ -repressor protein binds best to the E^d molecule and yet no T cells were isolated that recognize this pair. In searching for an answer to this paradox, the authors discovered that the hypervariable segment containing the internal motif shares three additional residues with the λ -repressor peptide; it does not share these with the other peptides binding to E^d. Based on this observation they speculate that these additional shared residues of the hypervariable segment might be exposed in such a way that they interact with the T-cell receptor in the absence of any foreign peptide and in the same manner that the epitope residues in the foreign peptide interact with the T-cell receptor

specific for the λ -repressor peptide (see closed state in the figure). Because these residues are part of a self-MHC protein, Guillet *et al.* postulate that the T cells capable of recognizing them (and also the λ -repressor peptide) must have been deleted during the induction of tolerance (non-responsiveness) to self proteins. Thus, despite the ability of the λ -repressor peptide to bind to the MHC molecule, an immune response cannot be elicited to this peptide because the T-cell receptors required to recognize it are no longer present in the repertoire.

In essence this new model divides the hypervariable segments of the MHC molecule into two subsites, an internal epitope that interacts with T-cell receptors during repertoire selection, and an internal agretope (motif) that interacts with the MHC molecule (see figure). The foreign peptide enters like a tape cassette when the player is open, binding to the MHC molecule via its agretope and displaying a new epitope to be read out by the T-cell population. In this model, alloreactivity (the ability of one T-cell population to recognize MHC molecules from another individual) is explained as recognition of the foreign MHC molecule in its closed position, with the residues of the internal epitope exposed for T-cell contact. These molecules are seen as foreign because they differ from self MHC molecules and hence have not had T cells reactive against them deleted by tolerance.

Support for this model comes from the recent work of Parham *et al.*¹¹ who showed that a peptide corresponding to one of the MHC hypervariable segments can inhibit the cytotoxic activity of an alloreactive T-cell clone. In addition, recent structural studies¹² on MHC class II molecules support the notion that internal interactions between MHC hypervariable segments are critical for the proper assembly of the MHC molecule. On the other hand, it is difficult to imagine how sharing a few amino acids between the primary sequence of a foreign peptide and the internal epitope would lead to an exact structural mimicry for T-cell recognition. Only X-ray crystallographic determination of the three-dimensional structure of the MHC molecule, with and without the foreign peptide, will clarify the situation. □

1. Reinherz, E. *Nature* **325**, 660-663 (1987).
2. Germain, R.N. *Nature* **322**, 687-689 (1986).
3. Schwartz, R.H. *A. Rev. Immun.* **3**, 239-263 (1985).
4. Babbitt, B. *et al. Nature* **317**, 359-361 (1985).
5. Rao, A., Faas, S. & Cantor, H. *Cell* **36**, 889-895 (1984).
6. Buus, S. *et al. Cell* **47**, 1071-1077 (1986).
7. Buus, S. *et al. Science* **235**, 1353-1358 (1987).
8. Guillet, J.-G. *et al. Nature* **324**, 260-262 (1986).
9. Guillet, J.-G. *et al. Science* **235**, 865-870 (1987).
10. Schwartz, R.H. *Nature* **317**, 284-285 (1985).
11. Parham, P. *et al. Nature* **325**, 625-628 (1987).
12. Braunstein, N.S. & Germain, R.N. *Proc. natn. Acad. Sci. U.S.A.* (in the press).

Ronald H. Schwartz is head of the Laboratory of Cellular and Molecular Immunology, National Institute of Allergy and Infectious Diseases, NIH, Bethesda, Maryland 20892, USA.

Meteorites

Diamonds are forever?

I.P. Wright and M.M. Grady

DIAMONDS were originally found in meteorites nearly 100 years ago in the achondrite Novo Urei (one of about 20 members of the ureilite group of meteorites). Other ureilites have since been shown to contain diamonds¹ and diamonds have also been unequivocally identified in the iron meteorites Canyon Diablo² and ALHA77283 (ref. 3). Although the existence of diamonds in meteorites is not questioned, their origin has at times generated considerable controversy. Urey⁴ chose a mechanism based on high static pressure, and postulated that the meteorites were formed in a lunar-sized body. Lipschutz and Anders⁵, on the other hand, proposed the currently favoured view that the diamonds were shock-produced, a mechanism not requiring inordinately large parent bodies. Two recent letters^{6,7} to *Nature*, however, show that not all extraterrestrial diamonds can be explained by this latter view. Diamonds have apparently been found in relatively unshocked carbonaceous and enstatite chondrites⁶; also lonsdaleite (hexagonal diamond) that is clearly not shock-produced has been found in interplanetary dust particles⁷. On the face of it, these results seem quite surprising; the stage now seems set for more controversies concerning the origin of meteoritic diamonds. A recent conference* provided a forum for discussion of the implications of diamonds found in primitive meteorites.

R.S. Lewis (University of Chicago) detailed how the diamonds were stumbled on during a search for the C₆₀ cluster called buckminsterfullerene. The Chicago group was looking for the host of an isotopically anomalous xenon component of meteorites characterized by an enrichment of light and heavy isotopes (Xe-HL, sometimes previously referred to as CCF-Xe). The product left after chemically concentrating the operationally defined host, known to be carbonaceous material, was a white powder. This result was odd as most carbonaceous residues from primitive meteorites are black. The Chicago group showed that the powder consisted not of C₆₀, but of fine-grained (~ 5 nm) diamonds. The group concluded⁶ that diamonds are the carriers of Xe-HL, although it should be stressed that, as yet, there is no demonstrable proof. One implication, if the diamonds are the host of Xe-HL, is that they were formed before the solar nebula collapsed to form the Sun and planets.

An important question this raises concerns the mechanism for xenon entrapment; Lewis proposes ion-implantation during a supernova explosion into existing diamonds. Depending on its kinetic energy, a xenon ion may pass through more than one diamond before it becomes implanted. (Note that only 1 in 10⁴-10⁶ diamonds will contain a xenon atom — most diamonds are xenon-free.) Another consideration is the carbon stable isotope composition of the diamonds. Doubtless astrophysical modelling will predict the ¹²C/¹³C ratio, but intuitively an enrichment of ¹³C might be expected if the diamonds had formed before the supernova occurred, during the red-giant phase of stellar evolution. But although stable-isotope measurements of the Xe-HL-host component show unprecedented ¹⁴N enrichments the carbon isotope composition is very similar to the average of the Solar System⁸.

Perhaps we should investigate the way in which diamonds could form in the solar nebula, rather than seeking some exotic origin. A recent News and Views article⁹ reported how diamonds can be made in the laboratory by various chemical vapour deposition (CVD) techniques under low pressures. Such reactions operating when the Solar System formed could have produced fine-grained diamonds. Indeed, this has been considered¹⁰ as an alternative origin for diamonds in ureilites that are commonly thought to be shock-produced. In the light of the recent discoveries we should reconsider the role of CVD in the formation of meteoritic diamonds.

Meanwhile, attention has been turned towards the isotope effects incurred during plasma-initiated gas-phase reactions, most notably a demonstration of mass-independent isotope fractionation of oxygen in ozone generation¹¹. M.H. Thieme (University of California, San Diego) found a similar isotope effect in reactions of CO. He also noted that carbon may be deposited as a by-product. It would be interesting to study the nature of this carbon, as large isotope enrichments of ¹³C can be produced by this reaction. Evidence for ¹³C-rich materials in meteorites abounds; E.K. Zinner (Washington University, St Louis) reported that although the carbon in hibonite grains in carbonaceous chondrites is isotopically normal, 22 per cent of spinel grains studied contain excess ¹³C. The nature of the association of carbon with these refractory minerals is still unclear, but in spinels, at least, it is considered that it may exist as silicon carbide or even

* Eighteenth Lunar and Planetary Science Conference Houston, Texas 16-20 March 1987.

diamond, and that these grains acted as seeds for mineral growth.

The study of carbon in interplanetary dust using microscopy and ion-probe analysis continues to produce interesting results. The link between cosmic dust, meteorites and even the solids detected from comet Halley is now being studied intensively. There have been no further observations of lonsdaleite, but undoubtedly they will be sought. Meanwhile, the shocked diamond-bearing ureilites also look set for re-examination, especially in the light of the speculation of C.A. Goodrich (University of Arizona) that they may be of venusian origin. □

1. Vdovynkin, G.P. *Space Sci. Rev.* **10**, 483–510 (1970).
2. Fronzel, C. & Marvin, U.B. *Nature* **214**, 587–589 (1967).
3. Clarke, R.S. Jr, Appleman, D.E. & Ross, D.R. *Nature* **291**, 396–398 (1981).
4. Urey, H.C. *Astrophys. J.* **124**, 623–637 (1956).
5. Lipschutz, M.E. & Anders, E. *Geochim. cosmochim. Acta* **24**, 83–105 (1961).
6. Lewis, R.S., Ming, T., Wacker, J.S., Anders, E. & Steel, E. *Nature* **326**, 160–162 (1987).
7. Rietmeijer, F.J.M. & Mackinnon, I.D.R. *Nature* **326**, 162–165 (1987).
8. Lewis, R.S., Anders, E., Wright, I.P., Norris, S.J. & Pillinger, C.T. *Nature* **305**, 767–771 (1983).
9. Roy, R. *Nature* **325**, 17–18 (1987).
10. Arrhenius, G. & Alfvén, H. *Earth planet. Sci. Lett.* **10**, 253–267 (1971).
11. Thiems, M.H. & Heidenreich, J.E. *Science* **219**, 1073–1075 (1983).

I.P. Wright and M.M. Grady are at the Department of Earth Sciences, Open University, Walton Hall, Milton Keynes MK7 6AA, UK

Regulation of transcription

Are some controlling factors more equal than others?

Nicholas J. Short

THE net is beginning to close on the shoal of proteins that dictate how well each gene is expressed in a eukaryotic cell. In particular, most promoters for RNA polymerase II so far examined contain binding sites for one or more soluble proteins (the so-called *trans*-acting factors) which increase transcription *in vivo* and *in vitro*. Recent work by Robert Tjian and colleagues reports¹ the purification of another such factor, but gives the story a novel twist by showing that the protein can perform functions other than elevating transcription levels. It also allows some important generalizations to be made about such proteins as a class.

The paradigmatic *trans*-acting factor is the protein Spl, first described as a protein necessary for the optimal functioning of the simian virus 40 (SV40) early promoter *in vitro*. Subsequent work in Tjian's laboratory and elsewhere culminated recently in the purification of the protein to near-homogeneity by a powerful new technique using affinity chromatography². Spl recognizes and binds specifically to a GC-rich decanucleotide sequence, which often appears in multiple copies upstream of the affected promoter. The stereochemistry of binding is important, as the multiple binding sites preceding the SV40 promoter are all on the same side of the DNA, and the efficacy of the sites in stimulating transcription depends strongly on how many helical turns of DNA separate them from other elements in the promoter (such as the enhancer or the so-called TATA box³). Inversion of the binding sites relative to the rest of the promoter does not, however, affect transcription. Although this property is also characteristic of eukaryotic enhancers, the effects of an Spl binding site differ sharply from those of an enhancer in that they fade

away over distances of a few tens of base pairs. The consequences of Spl binding are very similar *in vivo* and *in vitro*: the transcription rate increases 10–25 times, depending on the gene. Spl does not seem to be tissue-specific as many genes possess Spl binding sites and Spl is found in nuclear extracts from various tissues.

Several sequences that potentiate transcription by RNA polymerase II are recognized by factors resembling Spl. One such sequence is the CCAAT motif which occurs, among other places, upstream of various globin gene promoters, the herpesvirus thymidine kinase gene, and the *hsp 70* heat-shock gene. Nuclear extracts from cultured murine erythroleukaemia cells and HeLa cells turn out to contain a factor, CTF, that recognizes this sequence and stimulates transcription about 10-fold *in vivo* and *in vitro*. In the latest work from Tjian's laboratory, Jones *et al.* report⁴ the purification of CTF to apparent homogeneity using the same affinity technique with which they purified Spl. Even more strikingly, they also show beyond reasonable doubt that CTF is identical to a protein originally identified in a quite different context, nuclear factor I (NF-I).

NF-I was first described as a *trans*-acting factor required for the replication of adenovirus DNA. Binding of NF-I to a specific sequence at the adenovirus origin of replication is necessary to raise replication above very low levels *in vivo* and *in vitro*. This sequence, or some variant of it, is known to occur in several viral promoters and to be very similar to the known binding sites of CTF. That CTF and NF-I are identical was therefore not entirely unexpected. Proof of it, however, is welcome, as it demonstrates unequivocally that a single factor can influence both transcription and viral DNA replication,

and allows information obtained about it in either of its guises to be pooled. In particular, the binding site at the adenovirus origin of replication has an element of dyad symmetry, and shows a considerably higher affinity for the factor than most sequences recognized in promoters. Sites in some promoters, such as that in the human Harvey *ras-1* oncogene, share the high affinity and dyad symmetry; the affinity is raised further by mutations that make the adenovirus binding site perfectly symmetrical. Like Spl, NF-I/CTF seems to bind largely in the major groove on one side of the DNA⁴, and altering a symmetrical binding site in such a way that the two protein subunits can no longer rest on the same side of the DNA reduces the binding affinity. In addition, the exact separation between the NF-I/CTF binding site and the adjacent sequence where the viral terminal protein acts (another event required for viral DNA replication) has a marked effect on replication levels. Other points of resemblance to Spl are that the adenovirus binding site can be inverted without affecting replication; that effects of the factor on promoters only occur over relatively short distances; and that the factor is not tissue-specific.

Given the list of common features, it seems reasonable to suppose that NF-I/CTF and Spl act in a similar way to stimulate transcription. Indeed, what is known of other *trans*-acting factors suggests that most of them also exhibit some of these properties, and that NF-I/CTF and Spl are the best-characterized representatives of a new class of DNA-binding proteins with a common, as yet unknown, mode of action. It also seems plausible that NF-I/CTF stimulates viral DNA replication by this same mechanism. Although there is no hard evidence for it, this attractive hypothesis is supported by the number of distinct proteins that are involved in each task. Initiation of adenovirus DNA replication appears to require not only NF-I/CTF and the viral terminal protein, but also two more *trans*-acting factors, NF-III⁵ and the so-called octamer-binding protein, both of which have binding sites in many RNA polymerase II promoters. Similarly, NF-I/CTF binding sites occur in other such promoters alongside those for additional factors such as Spl. This may well be a general feature of binding sites for these proteins, reflecting either a combinatorial role for the different factors in controlling transcription, or some other unknown aspect of their action. Many known enhancer sequences, for instance, contain such clusters of binding sites next to other repeated motifs whose significance is not understood.

Perhaps the most interesting problems arising from the work concern how these factors act and why they show so little tissue specificity. The answer to the first question will undoubtedly become clearer

as more such factors are affinity-purified, cloned and eventually synthesized in bulk. The second may be less straightforward. So far, none of these factors has been shown to be very tissue specific or to increase transcription even 100-fold *in vivo* or *in vitro*, yet the expression of eukaryotic genes can vary from tissue to tissue over many orders of magnitude. Even the factor NF- κ B⁵, which binds to the immunoglobulin κ enhancer and seemed to be confined by κ -producing cells, now appears to be inducible in HeLa cells by phorbol ester treatment⁷ (although it differs in several ways from the Spl-type factors described here).

The possibility therefore remains that almost all the *trans*-acting factors uncovered in the remarkable discoveries of the

past three years constitute a mechanism for the fine-tuning of transcription in a given cell, and that the availability of a particular gene for transcription is determined by a separate class of *trans*-acting factors working in a different and unexplored way. □

1. Jones, K. A., Kadonga, J. T., Rosenfield, P. J., Kelly, T. J. & Tjian, R. *Cell* **48**, 79–89 (1987).
2. Briggs, M. R., Kadonga, J. T., Bell, S. P. & Tjian, R. *Science* **234**, 47–52 (1986).
3. Takahashi, K. *et al.* *Nature* **319**, 121–126 (1986).
4. de Vries, E., van Driel, W., van den Heuvel, S. J. L. & van der Vliet, P. C. *EMBO J.* **6**, 161–168 (1987).
5. Prijn, G. J. M., van Driel, W. & van der Vliet, P. C. *Nature* **322**, 656–659 (1986).
6. Sen, R. & Baltimore, D. *Cell* **46**, 705–716 (1986).
7. Sen, R. & Baltimore, D. *Cell* **47**, 921–928 (1986).

Nicholas J. Short is at the Department of Biophysics, King's College London (KQC), 26-29 Drury Lane, London WC2B 5RL, UK.

Drosophila genetics

Love-song and circadian rhythm

Michael Ashburner

BIOLOGICAL clocks allow organisms to synchronize their activities with the predictable cycle of day and night, giving them a sense of time, so that even in the absence of external cues an organism's behaviour changes during the day as it would under normal conditions. A population of the fruitfly *Drosophila*, for example, entrained to a 12-h light:12-h dark cycle, will continue to hatch just after dawn, and only then, even in constant darkness. At least six different genes are known in *D. melanogaster* to affect circadian rhythms. The best studied of these is the *per* gene, discovered by Ron Konopka when a graduate student at Caltech. New results reported by Yu and colleagues on page 765 of this issue¹ suggest a mechanism by which this gene mediates two different effects.

Mutation of *per* abolishes completely all circadian behaviour, not only the hatching rhythm but also the rhythm of spontaneous locomotor activity of adult flies. Other mutations of this gene can increase the period of the rhythm from 24 to 29 h (*per^s*) or decrease it to 19 h (*per^l*). Mutations of this gene also have remarkable effects on a rhythm with a very different periodicity — the oscillation in the interval between the bursts of sound that characterize the love-song male *Drosophila* sing to potential mates. This interval varies around a mean of 34 ms in *D. melanogaster* with a period of about 55 s in the wild type. In *per^s* males this period is 80 s, in *per^l* it is 40 s and in *per⁰* there is no rhythmicity in the variation in interpulse interval.

The temporal characteristics of the male love-song are highly species-specific. For example, in the closely related species *D. simulans* the mean interval between

pulses of song is 49 ms, fluctuating with a period of 40 s. Females are excited by a song characteristic of their own species, not that of another. More strikingly, hybrid females prefer the (simulated) song of hybrid males². This raises two questions. The first is the role of the *per* gene in the female; does it perhaps encode an oscillator against which incoming song is measured? If so, then this would neatly explain the genetic coupling between the male's stimulus and the female's response. The second is how the product of one gene can affect a species-constant circadian rhythm and a rhythm that is not only variable between species but whose very variation is concerned with species isolation.

Whether *per* encodes a component of the fundamental oscillator or a component of the black box that lies between the oscillator and the behaviour is not yet known. Perhaps molecular studies of *per* and its gene products will help^{3,4}. These studies are certainly providing fascinating insights into the genetic control of behaviour. As reported by Yu *et al.* in this issue¹ the effects of *per* on the very short period rhythms of the love-song and on circadian rhythms can be experimentally distinguished. Judging from its DNA sequence, *per* encodes a very large protein of 1,218 amino acids. These include a run of alternating glycine and threonine residues. The length of the Gly–Thr run varies between different wild-type *per* alleles: in some stocks it is 17 Gly–Thr repeats, in others 20 or 23. It is not known whether or not these different alleles are functionally different. We would certainly like to know whether males with these different alleles show the same periodicity characteristics of their love-songs. By some nifty cutting and splicing Yu *et al.*¹

constructed a *per* gene that lacks the DNA encoding the Gly–Thr repeat region and used it to transform *per⁰* flies. The surprising result is that this mutant construct can rescue the circadian rhythm phenotype to normality but that the flies have a very short love-song rhythm. The *per⁰* flies transformed with a normal *per* gene have a love-song periodicity of about 60 s; when transformed by the Gly–Thr deletion they have a periodicity of about 40 s. In principle, this result gives a mechanism by means of which the periodicity of the love-song can evolve without affecting the control of circadian rhythms.

The love-song of *per⁰* flies transformed with the deleted gene shows the same periodicity, about 40 s, as that of *per^s*, known to be a single amino-acid substitution about 100 residues amino-terminal to the Gly–Thr region^{5,6}. The periodicity of circadian rhythms in *per* transformants is negatively correlated with amount of *per* messenger RNA⁶, although periodicities shorter than 24 h have not yet been achieved by over-expression. It is possible to reconcile these observations in a single model if we assume that the circadian oscillation depends on the absolute amount of *per* gene product and that the song rhythm depends on variations in *per* protein concentration. As a first approximation it is easy to see how the amount of *per* protein could oscillate around a mean if, for example, the protein is unstable and blocked the translation of its own messenger RNA. Then a change in its stability (for example, by a change in the number of Gly–Thr repeats) will change the periodicity of oscillation of protein concentration. Such speculation is, however, probably premature. Although we know that the *per* protein is a proteoglycan-like molecule^{3,7}, there are also clear indications that the *per* gene encodes at least three different but related proteins⁴ that may be involved in different aspects of its function. Complementary DNAs encoding two of these proteins can rescue the circadian phenotype of *per⁰* flies. Moreover, there remains a large black box between *per* and its proteins and the behavioural phenotype and there is plenty of room for other genes and their products to interact with those of *per*. That having been said, we are but on the threshold of an exciting analysis of the genetic and molecular basis of a very fundamental behavioural pattern. □

1. Yu, Q. *et al.* *Nature* **326**, 765–769 (1987).
2. Kyriacou, C. P. & Hall, J. C. *Science* **232**, 494–497 (1986).
3. Jackson, F. R., Bargiello, T. A., Yun, S.-H. & Young, M. W. *Nature* **320**, 185–199 (1986).
4. Citri, Y. *et al.* *Nature* **326**, 42–47 (1987).
5. Yu, Q. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **84**, 784–788 (1987).
6. Baylies, M. K., Bargiello, T. A., Jackson, F. R. & Young, M. W. *Nature* **326**, 390–392 (1987).
7. Reddy, P., Jacquier, A. C., Abovich, N., Petersen, G. & Rosbash, M. *Cell* **46**, 53–61 (1986).

Michael Ashburner is at the Department of Genetics, University of Cambridge, Downing Street, Cambridge CB2 3EH, UK.

Cosmology

Superconducting cosmic strings

Craig Hogan

WITH all the recent talk of the Superconducting Supercollider, of supercomputers and supersymmetry and superstrings and cosmic strings, one is hardly surprised to hear of a new type of postulated object called a 'superconducting cosmic string'. This hybrid animal has recently joined the bulging bestiary of imagined fauna with which high-energy physics might populate the cosmos. Its formidable name is suggestive of several qualitatively new and dramatic effects it might have in various astrophysical environments as proposed in several recent papers, including that of A. Vilenkin and G. B. Field on page 772 of this issue¹.

Superconducting cosmic strings, like all cosmic strings, are large, astronomically extended structures — unlike the microscopic superstrings currently favoured as fundamental objects in the supersymmetric quantum theory of everything. (For example, see a recent News and Views article².) Let us recall, however, that cosmic strings are microscopically thin tubes of 'false vacuum', and within the string itself physics is very different from normal — the core of the string is composed of a piece of vacuum trapped in the highly excited state it had in the early Universe. In this nearly one-dimensional world, particles behave as they did in the first microsecond of the Big Bang: electrons and other fermions may become massless, and protons and neutrons may break up into their constituents, massless quarks. Of course, many aspects of physics in this domain are very uncertain, simply because we do not have experimental data to show how fields behave at very high energies or densities. For many aspects of string interactions, this uncertainty does not matter, because the gravitational effects of the strings are independent of the detailed microscopic physics going on inside. Most studies of astrophysical effects of cosmic strings — for example, galaxy formation, gravitational lensing, microwave background fluctuations and gravitational radiation background — rely only on those generic gravitational interactions that would occur for any type of cosmic string. But it is possible that the microphysics inside the strings could affect their macroscopic behaviour. In particular, Witten³ suggests that under certain theoretical conditions, strings might behave like superconducting wires in which there is no resistance to electrical current. This electromagnetic property is superimposed on the familiar interaction and behaviour of normal cosmic strings. For instance, they are still

expected to form in a cosmic phase transition, and oscillating string-loops of various astronomical sizes are still expected to be sprinkled about the cosmos.

Witten gives examples of two types of superconducting strings, fermionic and bosonic. In the fermionic string, a charged particle such as an electron becomes massless on the string. Thus it takes no energy to create electron-positron (e^-e^+) pairs on the string, and indeed they are trapped on the string because it would take an energy of $m_e c^2$ to remove a particle, as outside the string they are no longer massless. Charged particles thus travel along the string at the speed of light, carrying electrical current but encountering no resistance. This is true up to a certain maximum current; a limit is imposed because as the charge carrier density increases so does the maximum energy (or Fermi energy) of the electrons. If the Fermi energy approaches $m_e c^2$, the string starts to shed some of its charge carriers. In the bosonic superconducting string, the scalar fields which compose the string carry electromagnetic charge themselves, forming a frictionless Bose condensate similar to the Cooper-pair condensate in a laboratory superconductor. In this case, at some enormous current the string 'goes normal', losing its superconductivity suddenly and releasing a catastrophic burst of energetic particles.

Once again, the task of tangling with these unusual beasts is left to the astrophysicists. After a couple of years during which Witten's ideas have sunk in, people have noticed suddenly that we are surrounded by all sorts of superconducting strings in the Universe at large. I will describe briefly the examples that have come to my attention.

Chudnovsky and colleagues⁴ propose that a large (~ 30 pc) loop of cosmic string has actually been seen in the centre of our Galaxy. Images taken with the Very Large Array telescope in New Mexico at ~ 1 arc s resolution show thin threads which, to an optical astronomer, might look like fine hair fallen on the photographic emulsion. They suggest however that actually we are seeing synchrotron emission from relativistic particles in a bow shock created as the magnetic field wrapped around a superconducting string passes through the interstellar plasma. As the string is expected to move at more than a tenth of the speed of light, this hypothesis is easily tested, because the threads should move ~ 2 arc s yr^{-1} . This test may already have been carried out by the time this article comes to press. If it fails, the threads may

have to be explained using old-fashioned magnetic field-lines, as in filamentary structures found in solar flares.

To have a loop of string in the centre of our Galaxy is only likely if the mass parameter of the strings ($G\mu/c^2$, where μ is the mass per length) is less than $\sim 10^{-10}$, a factor of 10^4 less than the mass needed for strings which produce galaxies. Never fear, however, for superconductivity has its applications in more massive strings ($G\mu/c^2 \sim 10^{-6}$) too. Ostriker, Thompson and Witten⁵ suggest that superconductivity would radically change the way in which strings produce galaxies. If primordial magnetic fields had existed, they would have induced large currents (and hence magnetic fields) in string-loops as they formed. String oscillations would then produce low-frequency electromagnetic radiation, similar to that from the spinning magnetic field of the Crab pulsar. As in the Crab, these waves are absorbed by the surrounding plasma and heat it up. A typical oscillating loop can release an enormous ($\geq 10^{61}$ ergs) amount of energy in a relatively brief time, all of it dumped as heat in the intergalactic or pregalactic gas. It is argued that the Universe is populated by huge (≥ 20 Mpc diameter) expanding gas bubbles around these radiating loops. Galaxies are formed when bubbles collide, leaving behind the 'frothy' appearance of the present-day distribution of galaxies. Radiating loops are fairly rare in the present-day Universe, but given that this happened recently, perhaps there should be some observable side-effects — for example the formation of the cosmic X-ray and γ -ray background radiation. Because superconducting strings can have handedness (in the sense that the e^- and e^+ prefer to travel in opposite directions) there is even the possibility that they can generate the observed cosmic baryon asymmetry, the preponderance of matter over antimatter.

Yet more radical is the suggestion made by Vilenkin and Field in this issue¹, who propose superconducting strings as the central engine of quasars. The electromagnetic radiation from a string-loop, like its gravitational radiation, is not isotropic but is emitted in highly collimated beams from the cusps that tend to form as a string oscillates. The beaming is so strong that it can accelerate particles in the beam to ultrarelativistic velocities ($\gamma \geq 10^{14}$) and carve out a channel in the surrounding plasma, which might resemble the relativistic jets found in extragalactic radio sources. The favoured mass parameter for this model ($G\mu/c^2 \sim 10^{-7}$) is intermediate between the previous two schemes. This model must confront a wealth of observational data on quasars. It is remarkable that the gross energetics work out sensibly, and it will be remarkable indeed if detailed study of this idea uncovers no serious difficulty. As the evidence

for the currently favoured central engine — a massive rotating black hole — is admittedly circumstantial, it is stimulating to have a qualitatively different proposition for comparison. For example, a black-hole central engine would give oppositely directed jets separated by $\leq 10^{15}$ cm, but a string-loop emits jets from cusps on opposite sides, separated by $\geq 10^{21}$ cm.

Superconducting strings are too much fun to ignore, and so work continues on questions of current dissipation, interaction with plasma, magnetic-dynamo effects, back-reaction and recoil from the radiation beamed, and other physical properties. They are also useful as sources of cosmic rays (C. Hill, D. Schramm and T. Walker, Fermilab preprint) and of γ -ray bursts (A. Babul, B. Paczynski and D. Spergel, Princeton preprint), and probably of other interesting astrophysical effects as well. There is even some controversy about whether such ob-

jects are physically possible, or whether the intense magnetic fields near a string would create e^-e^+ pairs and rapidly dissipate current (P. Amsterdamski, personal communication). If superconducting strings are found, they will open up a whole new realm of fundamental physics; their properties convey specific information on the interactions of fields whose very existence is now only speculated. We will be fortunate indeed if nature is so obliging as to provide real-world examples of such a novel product of theorists' imaginations. □

1. Vilenkin, A. & Field, G.B. *Nature* **326**, 772 (1987).
2. Campbell, P. *Nature* **320**, 679 (1986).
3. Witten, E. *Nucl. Phys.* **B246**, 557 (1985).
4. Chudnovsky, E.M., Field, G.B., Spergel, D.N. & Vilenkin, A. *Phys. Rev.* **D34**, 944 (1986).
5. Ostriker, J., Thompson, C. & Witten, E. *Phys. Lett.* (in the press).

Craig Hogan is at the Steward Observatory, University of Arizona, Tucson, Arizona 85721, USA.

Protozoan physiology

Why are cows not green?

Ralph A. Lewin

ON page 790 of this issue, D.K. Stoecker and two colleagues at the Woods Hole Marine Biology Laboratory report some remarkable observations on the physiology and ecology of certain ciliated protozoa which contain chloroplasts and which occur in densities as high as three per millilitre of surface sea water. But protozoa are animals, you will protest: how do they come to have chloroplasts, which are normally parts of the cells of plants? And what are they doing with them? And, if this phenomenon is as common as Stoecker *et al.* indicate, how is it that it has been virtually unknown until now? I will try to explain.

Because you may not be familiar with the oligotrich ciliates in such genera as *Laboea*, *Tontonia* and *Strombidium*, I will begin with analogies drawn from much more familiar animals, cows. For our present purposes, the digestive system of a cow has a grinding apparatus at one end, a series of digesters and fermenters in the middle, and a waste-disposal pipe at the other end. Fresh, green grass, fed into the front end, is chewed up and transferred from one gastric hopper to another, where some of it is broken down by enzymes and the rest is expelled as methane and you-know-what. Although the biochemical breakdown is relatively complete, leading to acetic acid and other simple fuels, a few of the more valuable molecular constituents are retained and recycled more or less intact. Among these are the 10 or so essential amino acids that are needed in the diet of all animals (as far as I know) because their molecules are too complicated or

metabolically too expensive to synthesize. Much of the rest, carbohydrates, is converted and used as fuel to keep the cow warm and mobile.

Some marine molluscs do a rather better job at recycling components of their food. Having a delicately lobed digestive system, unlike that of the cow, they are

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Planktonic ciliate with sequestered chloroplast, viewed by transmission light microscopy (Photograph by Diane Stoecker).

able to conserve and reuse other dietary components. Some molluscs can accumulate complicated pigment molecules, which they store and can later eject when irritated. Others accumulate toxins from the seaweeds that they eat, and thereby in turn become toxic or nauseating to potential predators which therefore learn to leave them alone. Yet others store, intact, the stinging cells of coelenterates on which they browse, and use these to sting

creatures that try to eat them. And a few tectibranch sea-slugs have sufficiently elaborate digestive systems to be able to dismantle the protoplasm that they suck out of algae, digesting much of it but retaining the chloroplasts relatively intact. Not only intact but functional, so that some of these animals can photosynthesize, as it were vicariously, fixing CO_2 in the light and converting it to organic matter for their own nutrition.

Now, cows cannot do this: they are too clumsy and too opaque. But some otherwise colourless and very transparent protozoa can, and do, as Stoecker *et al.* show, and they can be quite efficient about it. These protozoa eat unicellular algae, digesting in their food vacuoles much of the cellular substance but leaving the chloroplasts intact and functional, and with these they can use light energy for photosynthesis. Are they really so much cleverer than cows? I think this is not the point; rather, that the ingested algae can (at least for a while) photosynthesize inside the ciliates as well as they did outside. If this process does not involve too much acid (which would degrade chlorophyll molecules) or proteases strong enough to destroy the structure of the chloroplasts, then the latter can go on working, making more oxygen and maybe also more food for the animals that contain them. Selection might then favour strains of the ciliates with even more discriminating digestive vacuoles, and that is perhaps how species of *Laboea*, *Tontonia* and *Strombidium* have 'learned' to keep ingested chloroplasts more or less intact and functional. In effect, these organelles enable the ciliates to eat less and to live relatively cheaply, like plants. (For more details, see the paper by Stoecker and her colleagues.) Sooner or later, perhaps, the adopted bits of plant cellular apparatus may wear out and have to be replaced. (But not in every case. In at least one genus of ciliates, *Mesodinium*, there seems to be a mechanism whereby the chloroplasts can be made to multiply, although this is not dealt with in the present paper.)

Biological recycling of this sort is apparently not uncommon. According to Stoecker *et al.*, ciliate cells with adopted chloroplasts occur with a frequency of 10^3 or more per litre of surface sea water, at least in the region of Woods Hole, Massachusetts, and presumably elsewhere. They might even make an appreciable contribution to the primary productivity of the area, though this has not been assessed here. Maybe the phenomenon has been generally overlooked because plankton samples tend to be pickled in formalin and stored for indefinite periods before they are examined, by which time the cells are all dead and mostly bleached. Now that the existence of chloroplastic ciliates has been recognized, and now that they can be grown in laboratory culture, a lot

more can be learned about them and about their physiology and ecology. And, now that their ecological importance has been recognized, a lot more will be.

If the work of Stoecker *et al.* had been published in 1887, someone would undoubtedly have pointed a moral to this interesting discovery. On the other hand, in 1987 we are more aware of the social needs for recycling various useful things

such as paper and rags, tyres and batteries from 'dead' cars, bottles and aluminium cans from drinking sprees. In any case, we certainly have a lot still to learn from the protozoa. □

Ralph A. Lewin is on sabbatical leave at the Botany School, Oxford University, from the Scripps Institution of Oceanography, University of California, San Diego, La Jolla, California 92093, USA.

Cosmology

Observing the Big Bang

B. E. J. Pagel

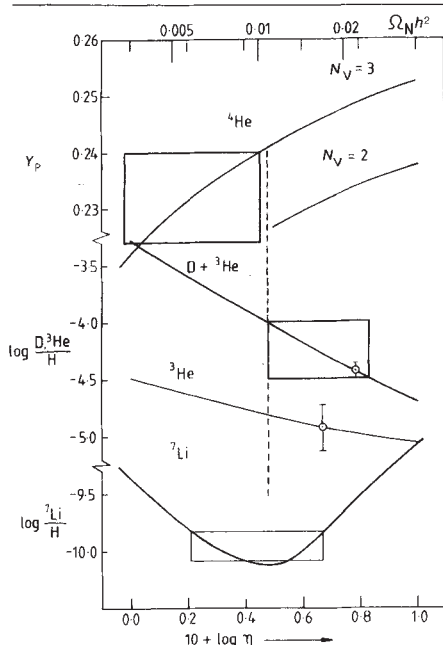
BIG-BANG cosmology has proved remarkably successful in explaining the cosmic microwave background radiation and it provides a framework for exciting new developments in high-energy physics. Primordial abundances of the light elements deuterium (D), ^3He , ^4He and ^7Li provide a critical test of the properties of the Universe during the epoch of nucleosynthesis between 1.5 and 5 min after the Big Bang, and they are in impressively fair accordance with the so-called standard model of a uniform expanding universe^{1,2}. But to account for the data this model requires an average density of baryons — particles such as protons and neutrons — at least 10 times less than the density of matter needed to 'close the Universe' in accordance with the Einstein-de Sitter model currently favoured by theorists, so that it only just fails to recollapse under its own gravity. This result has been alternatively regarded either as a convincing argument in favour of non-baryonic matter or as a weak point in the standard Big Bang nucleosynthesis (SBBN) model itself³. Density fluctuations that can in principle occur at various stages in the early Universe would modify the SBBN predictions³, and certain small discrepancies now being found between predictions and results derived from observation could suggest that such modifications are needed.

The figure shows the predictions of SBBN together with some data derived from observation. The predicted abundances depend on the ratio η of baryons to photons, which is related through the known temperature 2.7 K of the microwave background to the present density of baryons and h , the Hubble expansion parameter. The helium abundance depends in addition both on the number of species of light neutrinos, N_ν — normally taken as 3 for the three known leptons, the electron, muon and tauon — relativistic at the epoch of weak-interaction decoupling ($T \sim 1$ MeV), and on the half-life of the neutron.

In comparing theoretical predictions of primordial abundances with estimates of

the true values it is necessary to obtain the latter from various observations, coupled with an assortment of assumptions. In the cases of ^7Li , ^3He and D, prediction and observation appear to be in agreement. However, ^4He does not seem to fit the pattern so well.

For ^7Li the predicted abundance is uncertain by a factor of two. The observed abundance is based on results from old subdwarf stars and provides an excellent confirmation of SBBN predictions on the assumption that the ^7Li in subdwarf



A comparison of predicted and estimated primordial abundances of light elements. Predicted values are shown by solid lines; estimated values are shown by boxes, the error bars with 'Sun' symbols represent abundancies in the Solar System at the time of formation. Symbols: η is the ratio of baryons to photons; Ω_b is the fraction of baryon density to that required for closure of the Universe; h is Hubble's expansion parameter in $(10^{10} \text{ yr})^{-1}$. The predictions use a value for the half-life of the neutron of 10.4 min. The predicted value for ^7Li may be out by a factor of two. Note that if 1/4 of ^3He survives stellar processing (see text) then $10 + \log \eta$ must exceed 0.48 as shown by the broken vertical line.

atmospheres has suffered little depletion through mixing with hotter layers below. This assumption is supported by circumstantial evidence, but has no firm theoretical basis.

The observed deuterium abundance is based mainly on measurements of ^3He (partly a product of deuterium burning in the outer part of the Sun) in modern and ancient samples of the solar wind, and is in fair accordance with more direct measurements in planetary atmospheres and the interstellar medium. Because the effect of recycling interstellar gas through stars (astration) is always to destroy deuterium and never to create it, the observed abundance is a firm lower limit to the primordial abundance. But the correction factor for astration cannot be estimated with any reasonable precision because of uncertainties in models of the chemical evolution of the Galaxy. On the other hand, a proportion of ^3He survives astration, the proportion depending on the relative rates of formation of big stars and small stars — ^3He is destroyed in the former, but survives and could even be enhanced in the latter. It is estimated that overall at least 1/4 of the ^3He survives. This argument is used to deduce that a firm upper limit to the total abundance of primordial deuterium and ^3He relative to hydrogen^{1,2} is 10^{-4} by number of atoms.

The combination of this argument with SBBN and $N_\nu = 3$ imposes a strict requirement that the primordial ^4He abundance by mass fraction¹, Y_p , should exceed 0.24. The question of whether this requirement is met or not depends critically on the means of obtaining an experimental value. The most precise method of estimating Y_p is to observe recombination lines of hydrogen and helium in the emission-line spectra of bright ionized (HII) regions accompanying bursts of star formation in dwarf galaxies. The rate of stellar nucleosynthesis of helium is related to the production of other elements, such as oxygen and nitrogen. The primordial helium abundance, Y_p , can be deduced^{4,5} from plots of observed helium abundance versus observed abundance of such elements. A recent estimate⁵ from my group of primordial helium obtained in this way is $Y_p = 0.237 \pm 0.005$, a value consistent (just) with the magic figure of 0.24.

The above estimates were made using the assumption that the relevant He I atomic emission lines at 5,876, 4,471 and 6,678 Å are excited solely by recombination of He^+ ions with electrons, a process that is well understood⁶. In particular the excitation of He from the long-lived He 2^3S state to higher states was omitted. Recently, calculations of the rate coefficients for collisional excitation have been made⁷ for the lines at 5,876 and 6,678 Å. These have been used by Ferland⁸ in formulae that fit the observed line-strength ratios, but which imply an appreciable reduction

to previous estimates of helium abundance in H II galaxies.

Ferland has derived new values of Y_p based on measurements of the 5,876 Å intensities for several galaxies, incorporating quoted electron densities, and in three cases these are as low as 0.19–0.20. If correct, these values would raise grave doubts about the standard model. However, for easily identifiable reasons they are all either incorrect or doubtful.

Although he has exaggerated the drama of the situation, Ferland has performed a valuable service in drawing attention to the importance of the collisional contribution to helium lines in H II galaxies and has demonstrated that some downward correction to the deduced helium abundance is needed. Applying such a correction to our published data, we find $Y_p = 0.232 \pm 0.004$ (1 σ), very close to previous estimates by Peimbert⁹.

This value is 2 σ below the bare minimum requirements of SBBN with the limit on D + ³He, leaving three alternatives: (1) there are further systematic errors (not yet identified) in the determination of Y_p from H II galaxies; (2) the D + ³He argument is wrong because of unsafe assumptions about galactic chemical evolution, for example there could have been a

greater number of massive stars in the past than allowed for¹⁰ which would predict a large deuterium abundance in the Lyman- α forest clouds seen in absorption in the spectra of quasars, as these clouds are believed to represent material that has undergone relatively little processing since the Big Bang; (3) SBBN needs modification, possibly because $N_\nu = 2$ and not 3 (that is, the τ neutrino is massive) possibly because of fluctuations, or possibly for some other reason. Further developments are eagerly awaited. □

1. Yang, J., Turner, M. S., Steigman, G., Schramm, D. N. & Olive, K. A. *Astrophys. J.* **281**, 493–511 (1984).
2. Boesgaard, A. M. & Steigman, G. A. *Rev. Astr. Astrophys.* **23**, 319–378 (1985).
3. Sale, K. E. & Matthews, G. J. *Astrophys. J.* **309**, L1–L4 (1986).
4. Peimbert, M. & Torres-Peimbert, S. *Astrophys. J.* **193**, 327–333 (1974).
5. Pagel, B. E. J., Terlevich, R. & Melnick, J. *Publ. Astr. Soc. Pacific* **98**, 1005–1011 (1986).
6. Brocklehurst, M. *Mon. Not. R. astr. Soc.* **157**, 211–227 (1972).
7. Berrington, K. A., Burke, P. G., Freitas, L. & Kingston, A. E. *J. Phys.* **B18**, 4135–4147 (1985).
8. Ferland, G. J. *Astrophys. J.* **310**, L67–L70 (1986).
9. Peimbert, M. in *Star Forming Dwarf Galaxies* (eds Kunth, D., Thuan, T. X. & Tran Thanh Van, J.) 403–415 (Frontières, Paris, 1985).
10. Larson, R. B. *Mon. Not. R. astr. Soc.* **218**, 409–428 (1986).

B. E. J. Pagel is at the Royal Greenwich Observatory, Herstmonceux Castle, Hailsham, East Sussex BN27 1RP, UK.

Protein engineering

Inhibition into activation

Max Perutz

ON page 811 of this issue¹, Lau and Fersht report that directed mutagenesis of the enzyme phosphofructokinase (PFK) alters its properties in a manner not observed before in an allosteric protein, and apparently in conflict with allosteric theory. PFK, a tetrameric enzyme, catalyses the phosphorylation of fructose-6-phosphate (F6P) by ATP to form fructose-1,6-phosphate and ADP. The reaction is cooperative with respect to F6P and is inhibited by phosphoenolpyruvate (PEP).

X-ray analysis has shown that cooperativity is caused by an equilibrium between two alternative quaternary protein structures. Two catalytic sites of PFK are located at one pair of symmetry-related subunit interfaces; two regulatory sites occupy another such pair of interfaces². Precise knowledge of the stereochemistry at each of these sites has opened the way to the exploration of the catalytic mechanism by directed mutagenesis, in which single amino-acid residues can be replaced. Amino-acid replacements in the catalytic sites have provided proof that an aspartate close to the 1-OH of the bound F6P activates its nucleophilic attack on the γ -phosphate of ATP³. Directed mutagenesis in the regulatory sites, reported by Lau and Fersht in this issue¹, shows that

the replacement Glu 187 → Ala changes PEP from an inhibitor to an activator of the reaction of F6P with ATP, but at the same time reduces the maximal catalytic rate of which the enzyme is capable. How can this be reconciled with allosteric theory⁴?

According to this theory, proteins exhibiting cooperative behaviour are oligomers consisting of several subunits. These oligomers may exist in two or at least two different states that differ in their affinities for substrate or in their catalytic rates. Structurally, the states should differ in the arrangement of the subunits and the distribution and/or energy of the bonds between them, and therefore also in the constraints imposed on them. The stronger bonds constrain the protein to a tense (T) state with low substrate affinity or catalytic activity; the weaker bonds allow the protein to relax to the R state in which the subunits develop their full substrate affinity or catalytic activity. Allosteric effectors act by altering the equilibrium between the alternative states. In its mathematical form the theory can account for the properties of such proteins in terms of only three variables: the ligand equilibrium constants K_T and K_R (or $V_{\max}(T)$ and $V_{\max}(R)$), and $L = [T]/[R]$ in the absence of

substrate. The values of K and $V_{\max}$ are independent of the number of substrate molecules bound to the T or R states. Allosteric effectors would alter L , and be without influence on K_T and K_R . This formulation has led to Monod, Wyman and Changeux's theory being referred to (and frequently attacked) as the 'two-state model', but the text of their paper⁴ shows that they sensibly restricted their mathematical formulation to the simplest case, while being aware that the behaviour of real systems might be more complex.

Although allosteric theory was proposed more than 20 years ago, its predictions have been tested in detail only for the three allosteric proteins whose structures have been determined: haemoglobin, aspartate transcarbamylase and PFK. X-ray analysis confirms the existence of each of these oligomeric proteins in two alternative quaternary structures that differ in the arrangement of their subunits. The prediction of additional bonds between the subunits in the T structure has been confirmed in haemoglobin, but may not apply to the two enzymes, whose active sites actually lie at the subunit interfaces and must be maintained by bonds between them. Here only the distribution of the bonds may differ. In haemoglobin and PFK allosteric effectors are known to add to the constraints of the T structure by forming bonds between the subunits. These findings are in accord with the simple allosteric theory, but the thermodynamic behaviour of haemoglobin has turned out to be more complex. K_T varies over the equivalent of 2 kcal mol⁻¹ haem, depending on pH, [Cl⁻] and allosteric effector concentration⁵. Thus, K_T depends on the constraints of the T structure, which can therefore assume an infinity of thermodynamic states.

The mutant of PFK now discovered by Lau and Fersht is unique in switching the affinity of an allosteric effector from the T to the R structure. On the other hand, the reduction in the maximal catalytic rate is what one would expect if the binding of an allosteric effector to the R structure introduced bonds between the subunits which distorted the catalytic sites. In PFK, the active site lies between two domains of one subunit and one domain of a neighbouring subunit; catalytic activity may be influenced by rearrangements of the subunits as well as of the domains and also by changes in tertiary structure within the domains. X-ray analysis of the mutant enzyme should reveal which of these is dominant. □

1. Lau, F. T.-K. & Fersht, A. R. *Nature* **326**, 811–812 (1987).
2. Evans, P. R., Farrants, G. W. & Hudson, P. J. *Phil. Trans. R. Soc. B* **293**, 53–62 (1981).
3. Hellings, H. W. & Evans, P. R. *Nature* (in the press).
4. Monod, J., Wyman, J. & Changeux, J.-P. *J. molec. Biol.* **12**, 88–118 (1965).
5. Imai, K. *Allosteric Effects in Haemoglobin* (Cambridge University Press, 1982).

Max Perutz is at the MRC Laboratory of Molecular Biology, Cambridge CB2 2QH, UK.

Muscle contraction

Are two heads better than one?

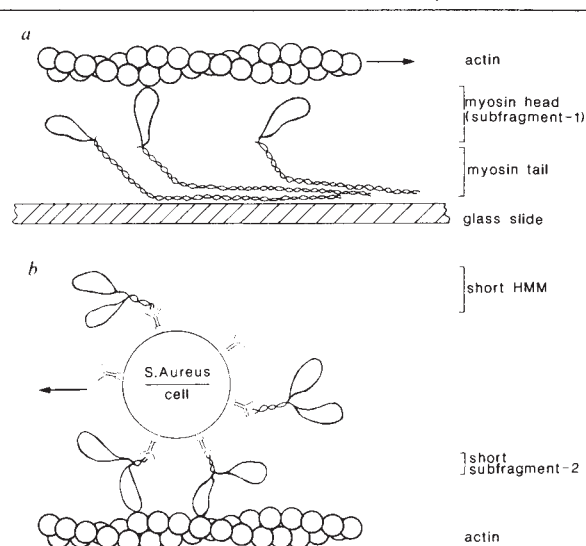
Clive R. Bagshaw

UNDER the electron microscope, the myosin molecule is seen to comprise two pear-shaped heads attached to a long tail. Two questions naturally spring to mind about the role of myosin in muscle contraction and other forms of motility. What is the significance of the two heads and where is the locus of force production? In the past, much of our understanding of the molecular basis of motility has come from the study of striated muscle, where the orderly arrays of filaments have aided the application of structural and physiological methods. Now headway is being made in less-structured preparations involving the direct observation of individual motile particles by light microscopy. Harada *et al.* on page 805 of this issue¹ show that single-headed myosin, prepared by controlled proteolysis, supports the motion of fluorescently labelled actin filaments; at the same time several lines of attack by Spudich's group^{2,3} are homing in on the conclusion that force can be produced by the interaction of the isolated myosin head (subfragment-1) with actin.

Most textbooks state, without evidence, that the two heavy chains of myosin, which make up the bulk of the head and the tail, are identical in structure. Even if this is true, the heads cannot simultaneously undergo identical interactions with a single actin filament, although the flexibility of the neck region may allow reasonable independence of their action. Nevertheless, one might speculate that the two heads must be there for a reason and possibly their cooperative action as they crawl, *tête-à-tête*, is fundamental to motion. The identity and interactions of the two subfragment-1 moieties, as judged by ATPase measurements, has remained controversial. Most groups admit to heterogeneity in their preparations; few have been able to demonstrate a 50:50 contribution as might be expected if the two heads were functioning in different modes. Addressing this question in structured systems is just as difficult. Using glycerinated muscle fibres, Sugi and colleagues⁴ recently attempted to test the necessity for two heads by correlating the extent of inhibition of ATPase activity by a thiol reagent, with the inhibition of tension generation. They concluded that a myosin molecule with one labelled head cannot generate tension, although its partner head displays its normal actin-

activated ATPase activity. The argument remains, however, a numbers game and, among other things, relies on the ATPase assay faithfully reporting the activity at high rates of ATP turnover in the fibre.

A more direct approach to the problem is to behead the myosin molecule and to examine the properties of the fragments produced. But what contractile properties can be measured in reconstituted systems?



a. The basis of the motility assay used by Harada *et al.*¹ to demonstrate that single-headed myosin filaments attached to a glass slide support the motion of fluorescent actin filaments. *b.* The *Nitella* system of Hynes *et al.*³ Short HMM was attached to *S. aureus* cells by protein A and an antibody against subfragment-2. Motion was observed along cables of actin filaments within a *Nitella* cell. Note that in *a* and *b* the direction of motion is controlled by the polarity of the actin filaments. Myosin heads that are in the wrong orientation cannot interact with actin and therefore do not give rise to an opposing force. (Components are not drawn to scale for clarity.)

Among the most promising are those that rely on looking directly at movement. Kron and Spudich² showed that myosin filaments spontaneously adhere to a glass slide and support the motion of actin filaments in the presence of ATP. They stabilized the actin filaments with a fluorescent derivative of phalloidin and under the microscope saw individual actin filaments squirming across the field of view, reminiscent of worms in an angler's bait tin. In this issue¹, Yanagida and colleagues now extend this approach to myosin filaments comprising single-headed myosin. They still observe movement when the single heads were diluted out with myosin rods so that the former comprised less than 10 per cent of the filament, indicating that single heads can operate independently (*a* in the figure). The velocity of movement (about $5 \mu\text{m s}^{-1}$) is similar for both single-

and double-headed myosin, but there is a different dependence on salt concentration. The affinity of myosin heads for actin is very dependent on ionic strength and in these assays there has to be a balance between maintaining contact of the filaments without seizing up their interaction.

In the meantime, Spudich's group³ has developed another assay system to identify the domains involved in force production. This group shows that short heavy meromyosin (HMM, a fragment containing two heads and a 40-nm tail) can drag an inert particle (actually a dead *Staphylococcus aureus* cell) along cables of actin filaments of a freshly dissected algal cell, *Nitella* (*b* in the figure). The significance of this finding is that short HMM lacks the hinge region in its tail that Harrington⁵ had proposed to be essential for force production through melting of its helical structure. It should be added, though, that Harrington honourably donated the antibody that allowed Spudich's group to attach the short HMM to the *Staphylococcus* cell.

Taken together, these observations suggest that an isolated head is capable of motion. This suggestion has just been confirmed. At the American Biophysical Society meeting in New Orleans (February 1987) Spudich reported the work of his colleagues, Toyoshima and Kron, who have succeeded in observing movement of actin filaments over subfragment-1 (generated by papain digestion) attached to a glass slide. In addition, De Lozanne's work was presented showing that a functional HMM can be expressed in the slime mould *Dictyostelium*, thereby opening up the prospect of more tests of structure-function relationships within the domains of myosin by protein engineering.

Although these novel movement assays report motion ($1-5 \mu\text{m s}^{-1}$) close to that of the relative sliding velocity of thick and thin filaments in a muscle sarcomere, the problem of quantifying force production remains. Calculations show that the filaments themselves, or the particles to which they are attached, offer very little resistance to flow and thus movement mimics unloaded shortening in muscle. Harada *et al.* in this issue¹ argue that filament movement at low concentrations of ATP indicates that single-headed myosin can generate nearly as much force as double-headed myosin, in order to overcome the resistance of other heads which are temporarily attached in a rigor complex, a state in which stiffness of the fibre is maximal. But it is not known whether actively contracting myosin heads can accelerate the dissociation of the rigor bond or whether they simply wait for the

latter to bind ATP. Regardless, the new movement assays will again focus attention on subfragment-1 as the site of force generation in muscle, rather than subfragment-2.

Those who have seen video tapes of the new experiments cannot fail to be impressed, but several groups will not be surprised by the outcome. Yano *et al.* reported⁶ that chymotryptic subfragment-1 (even smaller than the papain fragment used by Spudich's group) causes the slow movement of a rotor to which actin had been attached, presumably by generating streaming of the nearby solvent. In addition, a natural single-headed myosin has been known for some time to be a constituent of the protozoan *Acanthamoeba*. The short tail of this species is distinct from that of 'conventional' myosin, but it supports movement in the *Nitella* assay⁷. No doubt other primitive myosins are

waiting to be fished out with nucleotide probes or specific antibodies. As far as muscle is concerned, the α -helical coiled-coil tail is essential for the formation of the thick-filament backbone. Thus the facetious comment that 'myosin has two heads because it has got two tails' may not be too far from the truth. □

1. Harada, Y., Noguchi, A. & Yanagida, T. *Nature* **326**, 805-808 (1987).
2. Kron, S.J. & Spudich, J.A. *Proc. natn. Acad. Sci. U.S.A.* **83**, 6272-6276 (1986).
3. Hynes, T.R., Block, S.M., White, B.T. & Spudich, J.A. *Cell* **48**, 953-963 (1987).
4. Chaen, S., Shimada, M. & Sugi, H. *J. biol. Chem.* **261**, 13632-13636 (1986).
5. Harrington, W.F. *Proc. natn. Acad. Sci. U.S.A.* **76**, 5066-5070 (1979).
6. Yano, M., Yamamoto, Y. & Shimizu, H. *Nature* **299**, 557-559 (1982).
7. Albanesi, J.P. *et al. J. biol. Chem.* **260**, 8649-8652 (1985).

Clive R. Bagshaw is in the Department of Biochemistry, University of Leicester, Leicester LE1 7RH, UK.

Atmospheric physics

Magnetospheric forcing of upper atmosphere dynamics

Michael J. Rycroft

SINCE the beginning of the space age thirty years ago, *in situ* observations have been made by many different types of instruments aboard satellites orbiting the Earth. Direct measurements have revealed the tenuous, hot and weakly ionized upper atmosphere. Both the neutral thermospheric gas and the ionospheric plasma, at altitudes above 80 km, respond to ultraviolet rays and X rays from the Sun. It is only recently¹, however, that space physicists have realized the importance of the magnetosphere as a driver of upper atmospheric motions at high latitudes. This was the topic of a recent international symposium*.

Information on the electric field produced by the interaction of the solar wind with the magnetosphere is conveyed along geomagnetic field lines into the polar cap. There, and in the surrounding auroral oval, as illustrated in Fig. 1, a giant two-cell plasma convection pattern (in an inertial frame of reference) is usually created in the horizontal plane. Ions of the ionospheric plasma drag the neutrals along with them, by a frictional interaction. The neutrals are accelerated to velocities of several hundred metres per second, and heated by several hundred degrees. Field-aligned electric currents, considered by M. Blanc (Centre de Recherches en Physique de l'Environ-

nement Terrestre et Planetaire, France), transmit stresses from the boundary of the magnetosphere to the ionosphere, and vice versa. Joule dissipation of the million ampere currents flowing in the auroral electrojet (shown in Fig. 1) heats the upper atmosphere around the auroral

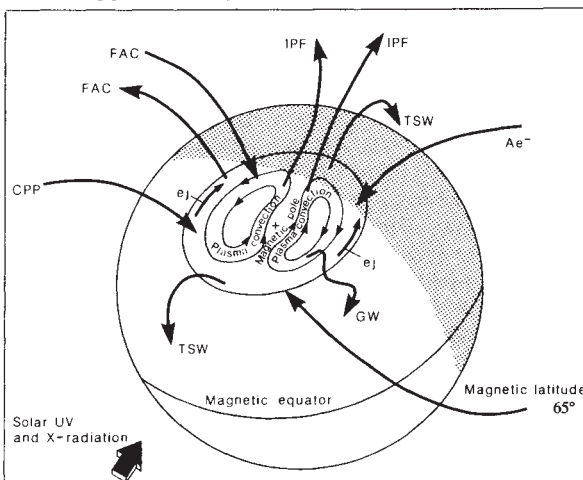


Fig. 1 Schematic view of the Earth's upper atmosphere, illustrating several regions and some of the physical phenomena that occur there. FAC, geomagnetic field-aligned currents; CPP, charged-particle precipitation from dayside cusp on magnetopause; IPF, ionospheric plasma flow up into magnetosphere; GW, gravity waves; TSW, thermospheric wind; Ae⁻, auroral electrons; ej, electrojet flowing along auroral oval.

oval. Spatially structured auroral electron beams that are accelerated to a few kiloelectron volts deposit their energy and cause the optical aurora, a subject reviewed by A.B. Christensen (Aerospace

Corporation, California). The electron beams also create structured secondary ionization in the auroral oval at altitudes above about 100 km. Such features evolve, via plasma instabilities, into irregularities with various scale sizes. Other irregularities are caused by structured electric fields.

Atmospheric buoyancy waves, commonly termed gravity waves, are launched by the heated auroral thermosphere and carry energy away to mid-latitudes. Although all these phenomena are usually studied in more detail in the north-polar regions, the high-latitude regions in the south are preferred for such investigations. This is because the distinction between solar and magnetospheric forcing of upper atmospheric phenomena is clearer in the Antarctic than in the Arctic as the geomagnetic pole is nearly twice as far from the geographical pole in the south as it is in the north.

M.A. Biondi (University of Pittsburgh), J.W. Meriwether Jr (University of Michigan) and V.B. Wickwar (SRI International, California) considered dynamical changes caused by the thermal expansion of the heated, high-latitude regions. Associated changes of ionic and neutral composition, their effects on electron recombination rates and hence on the ionospheric electron concentration, were presented by C. Lathulliere (Centre d'Etudes des Phenomenes Aleatoires et Geophysiques) and H. Rishbeth (University of Southampton). Not only satellite data are valuable here, but optical, radio and radar — particularly incoherent-scatter radar^{2,3} — observations from strategic places on Earth also contribute greatly.

A few groups are now working, both theoretically and experimentally, on the problem of the rapid (≤ 10 min) response of the ionosphere⁴, and the response of the thermosphere, smoothed over several hours, to the time-varying electric field at the magnetopause. The patterns of behaviour of the ionosphere and thermosphere are also being studied as long-term, large-scale averages. For example, Fig. 2 shows Southern Hemisphere satellite data on neutral thermospheric motions that have recently been analysed statistically to reveal the replacement of the standard twin-cell plasma convection pattern by a single vortex. A thermospheric wind blows strongly from noon to midnight over the geomagnetic pole, and blows westwards (towards earlier local times)

for several hours on either side of dusk, particularly at geomagnetic latitudes between 60° and 70°S. This single vortex is clearest when the dawn-to-dusk (east-west) component of the interplanetary

*Large Scale Processes in the Ionospheric-Thermospheric System. Boulder, Colorado, 2-5 December 1986.

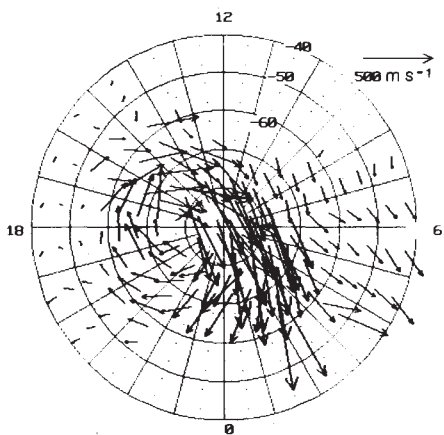


Fig. 2 Neutral wind vectors derived from data obtained by the low-altitude polar-orbiting Dynamics Explorer-2 (DE-2) satellite. The vectors show the influence of the east-west component (B_y) of the interplanetary magnetic field on the high latitude thermospheric circulation over the southern summer polar region. Data taken between November and January in 1981–82 and 1982–83, at all Universal Times, are plotted here in geomagnetic coordinates. The angle that the cross-polar-cap thermospheric wind vectors make with the noon–midnight meridian shifts from -7° for B_y positive to about 15° for B_y negative. (Courtesy of J.P. Thayer, T.L. Killeen and N.W. Spencer.)

magnetic field, B_y , is negative. When B_y is positive, as illustrated on the right-hand side of Fig. 2, the westwards circulation is restricted to the dusk sector of the polar cap, and a poorly defined dawn vortex develops. The average orientation of the thermospheric wind vector across the polar cap is seen also to depend on the sign of B_y . This topic has now been investigated using incoherent-scatter radar results^{5–7}.

The equations of motion of a parcel of air, under the action of pressure gradients, the Coriolis force, Lorentz forces acting on the charged particles, ion heating via ion–neutral collisions and other localized heat sources, can be solved using a super-computer. In such a thermospheric, global-circulation model⁸, realistic simulations are being achieved for both zonal and meridional winds. They indicate the strong geomagnetic control of thermospheric behaviour and its spatial variations at high latitudes. The latest model of the upper atmosphere is an interactively coupled ionospheric–thermospheric one; A.D. Richmond and R. Roble (National Center for Atmospheric Research, Colorado), D. Rees (University College London), R.W. Schunk and J.J. Sojka (Utah State University) all presented new results using such models. The coupling is generally stronger in summer than in winter, as a result of the larger plasma concentrations at high latitudes in summer, and of the solar energy input that supplements the charged particle energy input.

The sophistication of the coupled self-consistent ionospheric–thermospheric

Charlotte Friend (1921 – 1987)

CHARLOTTE Friend, who died on 13 January, loved working in the laboratory. She wanted only to have enough money and enough time to “play”, as she called it. This playing resulted in many discoveries, two of which have had a major impact on our understanding of cancer biology and tumour virology. In 1956 she discovered a virus-like agent that caused a malignant disease of the haematopoietic system in mice. She often recalled that the first time she presented the data at a meeting, there was great scepticism, in line with the climate of the time in which theories of the viral aetiology of cancer were not viewed favourably. She was not defeated by these negative reactions, and once joked that those who persisted in their heretical beliefs about tumour viruses were being accused of “having either holes in their heads or holes in their filters”. During this time, she found a great supporter in Peyton Rous who 40 years before had had to cope with similar reactions to his isolation of a chicken sarcoma virus; they developed and maintained a wonderful relationship until Rous’s death in 1970.

The original isolate of the virus that came to bear Friend’s name produced a leukaemia associated with anaemia in mice inoculated when newborn and maintained its pathogenic properties after adaptation to adult mice. With distribution of the virus stock to other laboratories and passage from host to host, virus strains that produced polycythaemia rather than anaemia were obtained. The observation that these strains contained another agent, a defective spleen focus-forming virus (SFFV), raised questions about the role of this second virus in the pathogenesis of the leukaemia. Friend and others expended enormous effort trying to resolve the ensuing controversy. She maintained that the original virus strain was competent and helper-independent and that SFFV was not essential for development of the erythroleukaemia that that strain induced. But she did acknowledge that variants might arise during *in vivo* passage through recombination with cellular elements or endogenous viruses. Some of her latest work

involved the use of molecular probes to compare different virus strains and to analyse how they might have arisen.

Within a short time after isolation of the virus, Friend showed that reticulum cell sarcomas developed when fragments of virus-induced leukaemic tissue were implanted subcutaneously into syngeneic mice, and that permanent cell lines could be established in culture from such tumours. These Friend erythroleukaemia (FEL) cell lines, even when cloned, consisted of mixed populations of undifferentiated cells and erythroid precursors; the cells continued to produce virus and were tumorigenic. The fact that a small percentage of the cells expressed erythroid functions led Friend to explore the long-considered possibility that the maturation block in malignancy could be removed. This led her to discover that FEL cells in culture could be stimulated to differentiate along the erythroid pathway by addition of dimethyl sulphoxide to the medium. This was one of the first clear demonstrations that expression of the malignant phenotype could be reversed, and Friend cells became a widely used model for analysing the regulation of gene expression in cell proliferation and differentiation and for detecting compounds capable of inducing or inhibiting differentiation. Most recently, Friend cells and their response to dimethyl sulphoxide and other modulators of differentiation have served as the inspiration, as well as the prototype, for evaluating the potential therapeutic effects of differentiation-inducing agents in human cancer.

Charlotte Friend was a woman of strong convictions and a fighter. She had no compunctions about defending ideas and causes in which she believed. She loved life — the theatre, music, books, art — and always found a way to surround herself with these pleasures. She was a fervent supporter of the women’s movement and will be remembered for the encouragement and help she gave others as well as for her seminal scientific contributions.

Leila Diamond

Leila Diamond is at the Wistar Institute, 36th Street at Spruce, Philadelphia 19104-4268, USA.

models used now is remarkable. Using these models, new physical insights are obtained of the forces that drive upper-atmosphere movements. This assists the interpretation of observations taken at different places and times. Predictions undoubtedly will soon be made; these will then put to the test the ingenuity of experimenters around the world. Conversely, extensive simultaneous observations made by the global network of incoherent-scatter radars are now being compared with the predictions of theoretical models, putting the models themselves to the test.

In ways such as these, we can improve our understanding of our outermost environment. □

1. Hays, P.B. *et al.* *J. geophys. Res.* **89**, 5597–5612 (1984).
2. Rishbeth, H. *Nature* **313**, 431–432 (1985).
3. Lockwood, M. *et al.* *Geophys. Res. Lett.* **14**, 111–114 (1987).
4. Rishbeth, H. *et al.* *Nature* **318**, 451–452 (1985).
5. Foster, J.C., Holt, J.M., Musgrove, R.G. & Evans, D.S. *Geophys. Res. Lett.* **13**, 656–659 (1986).
6. Kamide, Y. & Richmond, A.D. *Geophys. Res. Lett.* **14**, 158–159 (1987).
7. Foster, J.C. *Geophys. Res. Lett.* **14**, 160–161 (1987).
8. Rees, D. *Adv. Sp. Res.* **5**, No. 7, 215–228 (1985).

Michael J. Rycroft is head of the Atmospheric Sciences Division of the British Antarctic Survey, NERC, High Cross, Madingley Road, Cambridge CB3 0ET, UK.

Evolution east of Eden

SIR—"No names, no pack drill"¹. As in military service, in science criticism is safe only if nonspecific; the converse also holds: citation of alternative hypotheses must be accurate. The paper by Cann *et al.*² on mitochondrial DNA and human evolution misses in this regard, thereby blemishing the presentation of a methodology that otherwise holds every promise of providing an extremely valuable source of information on hominid evolution.

The authors' conclusion from restriction mapping is that all mitochondrial DNAs stem from an African woman who lived about 200,000 years ago, whereas other geographic areas were colonized more recently. They also outline an alternative view, that *Homo* has been present in Asia as well as Africa for at least one million years and that the transformation of archaic to anatomically modern humans occurred in parallel in different parts of the world.

Unfortunately, the reference³ which they cite on *Homo* being in Asia for at least one million years says something rather different, beginning "A number of separate lines of evidence indicate that all Asian hominids are less than one million years old [emphasis mine]", and concluding "...the entire record of *Homo erectus* in Asia may only span a period of approximately 600 thousand years (that is from 0.9–0.3 Mya)."

What of the references^{4,5} that Cann *et al.* cite supporting the idea that the transformation of archaic to anatomically modern humans occurred in parallel in different parts of the world? Coon⁴ did hold such a view, but Wolpoff *et al.*⁵ actually present a "multiregional evolution hypothesis" in which different hominid gene pools were established during initial occupation. Afterwards, "the pattern of regional variation was maintained throughout most of the Pleistocene by a balance between the local forces promoting homogeneity and regional distinction (selection, drift) and multidirectional gene flow". This evolutionary pattern is neither merely parallel nor traceable to a single centre save in the sense that all hominid lineages trace to a single source sometime in the past.

There are real uncertainties in estimates of rates of mitochondrial DNA evolution⁶, and influences of stochastic processes from population size and structure should be considered⁷. Until then there is insufficient cause to discard the first half million years or so of Asia's hominid-fossil record, particularly as at least some of the skeletal characters^{8–10} that are said by Cann *et al.* to make it "unlikely that Asian *erectus* was also ancestral to *Homo sapiens*" also occur in fossil hominid material from Europe (Arago, Petralona) and

Africa (Bodo). Of greater weight on the side of hominid evolutionary continuity in Asia back beyond 200,000 years ago is that other skeletal and dental traits marking living Asian populations (frontal keel, torus mandibularis, os epactale, shovel-shaped central and lateral upper incisors) also occurred in fossil remains sampled from populations dated to several hundred thousand years before the supposed African replacement 200,000 years ago¹¹. To have these features disappear via extinction and evolve anew in populations fresh from Africa would merely require parallel evolution of a different sort.

ROBERT B. ECKHARDT

Department of Anthropology and
Graduate Program in Genetics,
The Pennsylvania State University,
University Park,
Pennsylvania 16802, USA

1. Wintle, A. D. in *The Oxford Book of Military Anecdotes*, (ed. Hastings, M.) 394–398 (Oxford, 1985).
2. Cann, R. L., Stoneking, M. & Wilson, A. C. *Nature* **325**, 31–36 (1987).
3. Pope, G. G. *Proc. natn. Acad. Sci. U.S.A.* **80**, 4988–4992 (1983).
4. Coon, C. S. *The Origin of Races* (Knopf, New York, 1962).
5. Wolpoff, M. H., Wu, X. Z. & Thorne, A. G. in *Origins of Modern Humans: A World Survey of the Fossil Evidence* (eds Smith, F. H. & Spencer, F.) 411–484 (Liss, New York 1984).
6. Wainscoat, J. *Nature* **325**, 13 (1987).
7. Avise, J. C., Neigel, J. E. & Arnold, J. J. *molec. Evol.* **20**, 99–105 (1984).
8. Stringer, C. B. *Courier Forsch. Inst. Senckenberg* **69**, 131–143 (1984).
9. Stringer, C. B. *Courier Forsch. Inst. Senckenberg* **69**, 167–175 (1984).
10. Andrews, P. *New Scient.* **102**, 24–26 (1984).
11. Weidenreich, F. *Bull. geol. Soc. China* **XIX**, 64–75 (1939).

Of how great significance?

SIR—Bazan *et al.*¹ deny the significance of the similarities between the nucleotide sequences that code for the scrapie-associated PrP27–30 protein and human immunodeficiency virus (HIV) that we recently reported². They argue that other regions of HIV and cellular genes, such as the rat cytochrome P450c, show similarities that are as, or more, significant. But these similarities comprise unique regions of similarity, with an extensive number of gaps, whereas ours comprise four consecutive regions of similarity. Bazan *et al.* omit the fact that the probability of obtaining the four reported similarities between PrP27–30 and HIV by chance is less than 1 in 1,000 while that of obtaining those with cytochrome P450c or the other HIV regions is about 1 in 100 (ref. 3). It is worth noting that Bazan *et al.* use one algorithm, ALIGN, to compare the PrP27–30 sequence with HIV and another, FASTN, to compare PrP27–30 with other regions of HIV and other products. The scores, as well as the mean and standard deviation of their distributions, appear not to be comparable.

Braun and Gonda⁴ deny the significance of the similarities between the amino-acid

sequences of PrP27–30 and the carboxyl terminus of HIV. Unfortunately, these authors misinterpreted or obviated the basis of our report. Their arguments are based solely on protein sequence comparisons and fail to consider the significance of the similarities that exist at the nucleotide sequence level. Notwithstanding this omission, the *P* value for the global alignment between the scrapie-associated PrP27–30 protein and the carboxyl-end of the HIV endonuclease is 0.1 which is in fact significant, although marginally.

Both Bazan *et al.* and Braun and Gonda argue against the significance of the alignment between PrP27–30 and HIV by grouping HIV together with visna virus and equine infectious anaemia virus (EIAV) and showing that the similarity between visna and PrP27–30 or EIAV and PrP27–30 are lower at both the nucleotide and amino-acid sequence levels. This, however, does not deny a possible common evolutionary origin for all these sequences. Even if HIV, visna and EIAV belong to the same group of viruses, it is not expected that they will undergo similar evolutionary rates.

We appreciate the pinpointing of the human errors we incurred in our report during the transcription of the sequences in the final formatting of the figures. Nevertheless, none of these errors affect the significance of the comparisons or the statistical values thereof.

ROBERTO PATARCA

WILLIAM A. HASELTINE

TERESA WEBSTER

TEMPLE F. SMITH

Laboratory of Biochemical Pharmacology,
Molecular Biology Computer Research
Resource,
Dana-Farber Cancer Institute,
Harvard Medical School,
Boston, Massachusetts 02115, USA

1. Bazan, F. J., Fletterick, R. & Prusiner, S. B. *Nature* **325**, 581 (1987).
2. Haseltine, W. A. & Patarca, R. *Nature* **325**, 115 (1986).
3. Smith, T. F., Waterman, M. S. & Burke, C. *Nucleic Acids Res.* **13**, 645 (1985).
4. Braun, M. J. & Gonda, M. A. *Nature* **325**, 113 (1987).

Alzheimer amyloid aspects

SIR—Kang *et al.* (*Nature* **325**, 733–736; 1987) suggest that the major protein subunit (A4) of the amyloid fibril of tangles and plaques of Alzheimer's disease is derived by inappropriate processing from a 695 amino-acid precursor polypeptide. As an alternative, we would like to suggest that there is inappropriate initiation of translation at methionine 596, which immediately precedes the A4 sequence.

Kozak (*Cell* **44**, 293–292; 1986) has described an optimal sequence for initiation of translation by eukaryotic ribosomes as ACCATGG, where at position

−3 (relative to the A of ATG) A is better than G which is much better than Y, and at position +4, G is better than T. These are the two most important positions. She has also shown that some initiation can occur at a 'better' sequence downstream from the first initiation codon. The nucleotide sequence surrounding methionine 596 is AAGATGG. Thus there is an A at −3 and a G at +4, which by Kozak's hypothesis makes the sequence a good initiation region. At the start codon predicted by Kang *et al.* the sequence GCGATGC has a G at −3 and a C at +4.

Of the 19 other methionine residues in the sequence only that at amino acid 141 is close to the optimal Kozak sequence. Indeed, less than 5 per cent of eukaryotic mRNAs have the ideal sequence suggesting that correct initiation is regulated by other factors. Thus it is conceivable that failure of regulation leads to inappropriate initiation at methionine 596 in Alzheimer's disease. The peptide formed would not necessarily be inserted into the cell membrane, overcoming the difficulty of intramembrane cleavage raised by Kang *et al.*, as that region of the peptide would be exposed in the cytoplasm.

LARS H. BREIMER

PAUL DENNY

Department of Chemical Pathology,
Royal Postgraduate Medical School,
Hammersmith Hospital,
London W12 0HS, UK

Chloranthoid angiosperms

SIR—The report by Friis *et al.*¹ providing new evidence for chloranthoid angiosperms based on fossil androecia with pollen *in situ* from the Lower and Upper Cretaceous, is a very stimulating addition to previous palaeopalynological knowledge, particularly that due to Muller² who succeeded in reconstituting the continuous history of the *Clavatipollenites-Ascarina* complex (Chloranthaceae) from the Aptian onwards. The palaeobotanical record shows that some chloranthoid angiosperms with small, simple, anemophilous or entomophilous, unisexual or bisexual flowers must have played a dominant part in the earliest phases of angiosperm evolution. Friis *et al.* agree that these results contrast with those of most neobotanists who consider the chloranthaceae flower derived, compared to the large Magnoliales chlamydous flowers. Only a few botanists^{3,4} have rejected this last view.

I would like to point out that as early as 1981 I published a paper⁵ entitled "An ancestral dicotyledon with a strobiloid flower, *Hedyosmum*, is living today". The male flower of this genus of Chloranthaceae, generally considered as highly advanced, was conceived as a primitive, primarily anemophilous, polystaminate strobilous (comparable to the male cone

in conifers), and derived from a spike of small and naked flowers not fundamentally unlike those of the extant *Ascarina*^{6,8}. I think the *Hedyosmum* female flower might be derived in the same way from a spike of inconspicuous flowers. Thus I emphasize in a subsequent paper⁶ that the new interpretation of the *Hedyosmum* flower leads me "to think very seriously that some (many?) angiosperm trends must have had a chloranthoid origin". The discoveries of Friis *et al.* are a strong new argument in favour of my views.

J.-F. LEROY

Museum National d'Histoire Naturelle,
16 Rue Buffon, F75005 Paris

1. Friis, E.M., Crane, P.R. & Pederson, K. R. *Nature* **320**, 163-164 (1986).
2. Muller, J. *Bot. Rev.* **47**, 1-142 (1978).
3. Burger, W.C. *Bot. Rev.* **42**, 345 (1978).
4. Meeuse, A.D.J. *All About Angiosperms* 39,126 (Eburon, Delft, 1987).
5. Leroy, J.-F. *Abstr XIII Int Bot. Congr.* Sydney, Australia, 136 (1981).
6. Leroy, J.-F. *Taxon* **32** 169-175 (1983).
7. Leroy, J.-F. *C. heb. Séance. Acad. Sci., Paris* **296**, 747-752 (1983).
8. Leroy, J.-F. in *Précis de Botanique 11, végétaux supérieurs* (eds. Gaussen, H., Leroy, J.-F. & Ozenda, P.) 200-473 (Masson, Paris, 1982).

Bacterial sex timely

SIR—In a Commentary¹ entitled "Postmature scientific discovery?", Zuckerman and Lederberg suggest that the discovery of sexuality in bacteria by Lederberg and Tatum in 1946 was postmature. Among the criteria they use to characterize such discoveries are that they could have been made earlier with then available methods, and that they answer previously unrecognized problems.

Could bacterial sex have been discovered earlier? The reminiscence and the historical record say no. Lederberg states that he had to develop doubly auxotrophic mutants of *Escherichia coli* specifically to search for sex once he discovered that the spontaneous reversion rate of *Neurospora* was too high to detect recombination. At least two approximately contemporaneous papers report searches for bacterial sex^{2,3}. The results were negative, because only single auxotrophs were used and insufficient numbers of bacteria were screened.

Was bacterial sex a legitimate research problem and, if not, why did Lederberg decide to search for it? In both ref. 1 and the accompanying reminiscence⁴, Lederberg states that the identification of DNA as the substance responsible for the transformation of *Pneumococcus* by Avery *et al.* was critical in triggering his investigations. "Questions about the biology of transformation would remain inaccessible to conventional genetic analysis if bacteria lacked a sexual stage" (ref. 4). Thus the work of Avery *et al.* made the question of bacterial sex "newly consequential." (ref. 2). The link between the discovery of change in genotype caused by a molecule

isolated from cells and the postulation of a bacterial sexual cycle needs to be clarified. How would sex illuminate the mechanism of transformation? The discovery of transformation gave Lederberg the hope that "the chemistry of the gene" could be understood. One could postulate that as the transforming principle was isolated from bacteria, they were appropriate organisms in which to study genes. By all precedents, the study of genetics required a sexual stage and therefore one had to establish sexuality in bacteria before advancing to the chemistry. Was this in fact Lederberg's reasoning?

Although the authors present a convincing account of the conventional wisdom's working against the study of sex in bacteria, there are indications that it was a legitimate research problem in the 1930s and 1940s.

The presence of bacterial sex was a problem which other investigators deemed important^{2,3}. The reaction of at least some of the scientific community to the data presented by Lederberg and Tatum, proposing ways to eliminate a possible alternative explanation for the result, also suggests that sex in bacteria was a legitimate problem which they wanted solved.

The unique contribution of Lederberg was his experimental design, the use of different triple auxotrophs in the input strains, so that the presence of cells prototrophic for all markers or combinations of markers in the absence of recombination would be the multiple of the reversion rate for each of the auxotrophic genes. This design is critically dependent on the work of Luria and Delbruck⁵, which was not published until 1943. There is therefore a very small window of time for which the Lederberg-Tatum experiments could be considered postmature."

Finally, if the authors wish to convince us that this observation could have made earlier, I believe they are obligated to identify the investigators who had the tools, both methodological and intellectual, to do so.

MURIEL LEDERMAN

Department of Biology,
Virginia Polytechnic Institute and
State University,
Blacksburg, Virginia 24061, USA

1. Zuckerman, H. & Lederberg, J. *Nature* **364**, 629-631 (1986).
2. Sherman, J. M. & Wing, H. U. *J. Bact.* **33**, 315-321 (1937).
3. Gowen, J. W. & Lincoln, R. E. *J. Bact.* **44**, 551-554 (1942).
4. Lederberg, J. *Nature* **364**, 627-628 (1986).
5. Luria, S. & Delbruck, M. *Genetics* **28**, 491-511 (1943).

Scientific Correspondence

Scientific Correspondence is intended to provide a forum in which readers may raise points of a scientific character. They need not arise out of anything published in *Nature*. In any case, priority will be given to letters of less than 500 words and five references. □

The explosive start

William H. McNeill

Science and Civilisation in China, Vol. 5: Part 7. The Gunpowder Epic.
By Joseph Needham. Cambridge University Press: 1987. Pp. 703. £50, \$99.50.

NO WORK of scholarship in the twentieth century has done as much to alter received ideas about the past as Joseph Needham's *Science and Civilisation in China*. This latest volume, *The Gunpowder Epic*, is indeed an epic. Like some of its predecessors, it far transcends the limits implicit in the title by dealing also with Moslem and Christian uses of gunpowder, mainly for war but also in peace.

Needham shows more fully than ever before, and with a richness and range of scholarship others have never attained, that Chinese Taoist alchemists were the first to explore the incendiary and explosive properties of gunpowder; and that Chinese military men pioneered a variety of uses for the mixture, up to and including guns and rockets, long before gunpowder reached western Asia and Europe. Yet, as he is at pains to point out, by 1500 European guns surpassed Chinese models, whereupon the Chinese, along with other peoples of the Earth, responded by imitating such things as detachable breech blocks for artillery, and first matchlock and then flintlock muskets.

The most important part of the story for us, however, lies in its beginnings among the Chinese. Until Needham and his assistants came along, neither Chinese nor European scholars had explored the vast array of texts in which Chinese alchemical and military engineering achievements were recorded. The obstacles were, and remain, formidable, in particular the problem of deciphering what the technical terms, which were often fanciful and sometimes deliberately archaic, actually refer to. Moreover, it was often state policy to keep new weapons and manufacturing methods secret, further confusing the record. Fortunately, contemporaries faced similar difficulties in making technical matters clear to their readers. They therefore resorted to drawings, many of which Needham reproduces. Indeed, one can follow the argument simply by looking at each of the 235 illustrations in the book, and perusing the captions underneath. Without these visual aids, neither the original texts Needham used nor his history would be intelligible.

Needham's talents are extraordinary — a combination of linguistic ability, chemical and technical competence, and a cast of mind that has put endless details together into a clear and convincing picture of a world-wide development that ran across some 1,500 years. In this volume he finds the roots of the discovery of gun-

sary to withstand the greatly increased pressure of such explosions, as attested by the earliest surviving Chinese cast-metal gun, which dates from 1288. But this chance find was clearly not the earliest weapon of this sort — Buddhist carvings in a cave in Szechuan, discovered in 1985 by one of Needham's assistants, Robin Yates, show a hand gun in action, and probably date from the mid-thirteenth century or soon thereafter (pp. 290 and 580–581).

Transmission of the technology to Europe and the Moslem world was rapid, for in 1326 a European manuscript illustrated a vase-shaped gun in action, much resembling the hand gun depicted in the

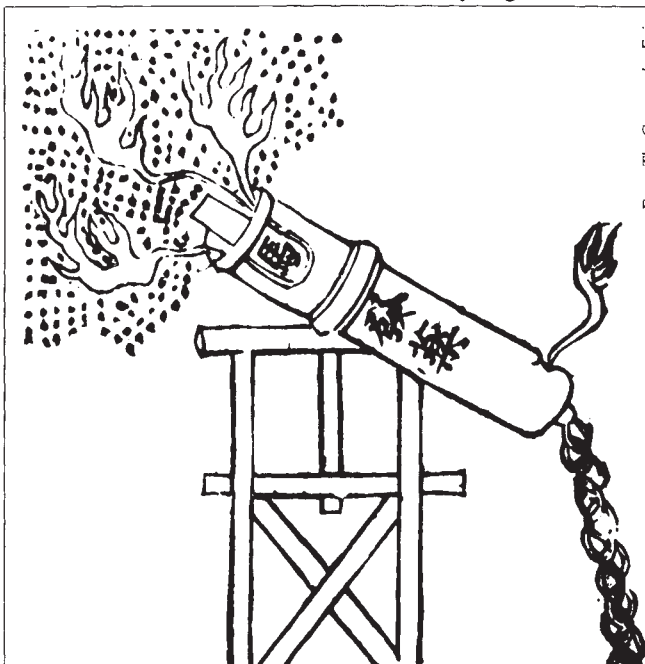
Szechuan cave. When, in due season, European gunmakers took the lead, the reverse flow of technical know-how was no less rapid, for a professional military establishment could never neglect a new weapon that promised superior performance — and the Chinese had possessed such an establishment for centuries before and after the development of the gunpowder weapons which they had pioneered.

Needham is quite as much interested in the peaceable uses of gunpowder as in its military applications. The scope and buoyancy of his inquiry is best revealed by the closing paragraph of his prefatory note:

So now let us pull the lanyard and fire off this gunpowder volume... upon the Republic of Learning, not indeed with the intention of doing any damage, but rather hoping that it may help those still looking for enlightenment about the history of gunpowder weapons and heat engines. War may or may not have been a decisive factor in human evolution and social progress, but what cannot be denied is that the steam engine and the internal combustion engine have been this, and all were children of the cannon. And that, in turn, was one development of the fire lance, while the other was the rocket, on which all space travel depends. Gunpowder engines and steam engines no less than the rocket vehicle were thoughts springing from the European Scientific Revolution — but all the previous developments, through eight preceding centuries, had been Chinese.

To have lifted the veil from those eight preceding centuries and cast the whole story into an epic frame is Needham's achievement. This is a truly great and mind-enlarging book. □

William H. McNeill is a Professor in the Department of History, University of Chicago, 1126 E. 59 Street, Chicago, Illinois 60637, USA. Among other books he is author of *The Pursuit of Power: Technology, Armed Force, and Society Since AD 1000* (Basil Blackwell, 1983).



The "heaven-rumbling thunderclap fierce fire eruptor", depicted in a Chinese military work of the early seventeenth century. The contraption fired a bomb or shell containing poisonous smoke.

powder in Taoist efforts to create life-prolonging elixirs. The effort began long before the Christian era, but experimenters clearly hit upon saltpetre and methods for its purification towards the close of the fifth century AD (p. 97). Needham then shows how Chinese military men used mixtures of saltpetre with sulphur and carbonaceous matter as an incendiary material, devising many ingenious ways of directing it against an enemy. For example, lengths of bamboo, filled with incendiary gunpowder and affixed to lances, produced a formidable 'five-minute' flame thrower; and when mixtures with a higher proportion of saltpetre were found to burn with explosive force, new possibilities opened up. Weapons that spat out missiles amidst flaming gases eventually became guns, after it was found that a fully occluded opening accelerated the projectile and increased its power enormously. Cast-metal tubes were, of course, neces-

From *The Gunpowder Epic*

In Epstein's image

C.T. Pillinger & I.P. Wright

Stable Isotopes in High Temperature Geological Processes. Edited by J.W. Valley, H.P. Taylor, Jr, and J.R. O'Neil. *Mineralogical Society of America: 1986. Pp.570. Pbk \$18.*

FOR years there has been a need for an advanced textbook on the geochemistry of stable isotopes (H, C, N, O, S and Si). This volume goes half way towards filling the gap by dealing with all aspects of stable isotopes in high-temperature geological processes. Given that its publishers have commissioned a companion publication to cover the subject from the low-temperature standpoint, the problem would now appear to be completely solved.

The book is divided into 14 chapters in which a multiplicity of experts cover theoretical aspects, igneous and metamorphic rocks, ore deposits, the characterization and effects of water-rock interactions, magmatic volatiles and meteorites. Each contribution can be read independently to satisfy an individual need and this is undoubtedly how the book will be used by most of its readers. Although it is priced at a level to be affordable by all, only the very brave, the dedicated or someone with an extended lecture course to give will attempt to read it from cover to cover. This is a pity, because the daunting mathematics of the first few chapters contain the quintessential basis of stable isotope geochemistry. The physical and chemical properties which allow stable isotopes to be partitioned between different phases or substances, and the rules which govern equilibrium and kinetic isotopic fractionation and isotopic exchange in open and closed systems, are the first principles of the subject. But be warned, these are 'first principles' for those who already know what stable isotope geochemistry is.

Elsewhere, the book is extremely readable. For instance, the chapters dealing with oxygen and hydrogen isotope compositions and their relationship with other systems such as Sr, Nd and Pb include a general introduction and overview, examples from the circum-Pacific regions, and case studies from Africa, Eurasia and Oceanic islands. This format will undoubtedly be appreciated by igneous geochemists who are prolific users of stable isotope data.

Two chapters address the subject of isotopic variations within the mantle. Of special interest is the origin and evolution of compositions observed for volatile species. This is a particularly active and controversial topic at present, and even though the relevant chapters were up to date in November 1986 there have already been additional papers to contribute to

the perspective. A couple of chapters likewise are devoted to problems involving the passage of aqueous fluids — the section on the isotopic characteristics and definitions of different types of natural waters is very useful.

The book is dedicated to Professor Sam Epstein in recognition of his contribution to stable isotope geochemistry over 40 years. Contribution is clearly not a strong enough word — many of the authors, now world leaders in the separate subject areas covered, owe their start directly or indirectly to Epstein's guidance and en-

couragement. It is a pleasure to join them in their tribute by recommending this book. Perhaps the extent to which the authors have succeeded in their aim is encapsulated in a sentence in the final section of the last chapter. It warns that "[stable] isotopic data alone cannot, under any circumstances, provide meaningful answers to any geologic problem". Such a cautionary note could imply that the gospel has been preached too well. □

C.T. Pillinger and I.P. Wright are in the Department of Earth Sciences, The Open University, Walton Hall, Milton Keynes MK7 6AA, UK.

Complementing the library

Kenneth Reid

Immunobiology of the Complement System: An Introduction for Research and Clinical Medicine. Edited by Gordon D. Ross. *Academic:1986. Pp.273. \$45.50, £38.*

THERE must be many people who feel overwhelmed by the current explosion in protein and DNA sequences, and in the uses of DNA probes that are being reported from all areas of biochemistry, including the complement system. They will be relieved to find that this multi-author volume, describing the many complement components, regulatory proteins and receptors, contains only a couple of short discussions of sequence data (the sections on the thiol ester sites in components C3 and C4 and the functionally important residues in the anaphylatoxins), while the review of the recent DNA work is quite brief. This feature is probably a strength rather than a weakness in an introductory account of this sort.

The opening section is a historical review, and contains eight pages of photographs of leading personalities from the early days of complement research. It starts with Nuttall's discovery in 1888 that normal sheep blood contains a heat-labile antibactericidal agent, and proceeds up to the early 1970s. This account will perhaps be of most interest to those already familiar with the system, and it is the following five chapters that make the book valuable as an introduction.

These contributions, which occupy about half of the text, cover the broader molecular aspects at a level which will enable the uninitiated to get a good grasp of the topic. They provide a clear overall view of the essential features of the activation of the classical and alternative pathways, the assembly of the terminal components, and the distribution and function of the increasing number of well-defined

receptors. Current views on the activation and control of the alternative pathway, which are often difficult for newcomers to understand, are particularly well described.

The remaining six chapters are of a more variable nature. But they illustrate clearly that any authoritative volume on complement now has to draw upon a wide variety of disciplines, including molecular biology, cell biology, pathology, genetics and various aspects of medicine. This latter part of the book provides a helpful survey of the role of complement in bacterial and viral infections, rheumatic diseases and red cell and platelet injury, as well as overviews of the genetics and biosynthesis of the complement components.

The editor has done well to eliminate repetitive descriptions of the activation of the system which could easily have been a problem in a volume of this type. The book's size and format are appropriate, and most chapters provide pertinent, up-to-date references (to 1985 in most cases) for further general reading. Details of the isolation or assay of any particular component were, of course, outside the scope of the book, but it is unfortunate that the authors of the early chapters did not at least provide some references to the literature on practical aspects.

Dr Ross has achieved his aim of producing a palatable introduction to the complement system for those interested in any of its structural, functional, genetic or clinical aspects. The book should therefore make a useful addition to the immunology section of most libraries. □

Kenneth Reid is Director of the MRC Immunochimistry Unit in the Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, UK.

New editions

• *Monoclonal Antibodies: Principles and Practice* 2nd Edn, by James W. Goding. Publisher is Academic, price is £17.50, \$29.50. For review of the previous edition see *Nature* **312**, 82 (1984).

• *Oxford Textbook of Medicine Vols I and II* 2nd Edn, edited by D.J. Weatherall et al. Publisher is Oxford University Press, price is £75 (£95 after 30 June), \$135. For review see *Nature* **303**, 204 (1983).

Design for living

J. R. Cann £28

How to Build a Habitable Planet. By Wallace S. Broecker. Eldigio Press, Box No. 2, Palisades, New York 10964: 1986. Pp.291. Hbk \$30; pbk \$18.

In *Nature* of 12 March 1987 Wallace Broecker described how he set about desk-top publishing, from the Lamont-Doherty Geological Observatory (Eldigio Press — got it?), to try to avoid what he sees as the stifling and unhelpful miasma of mainstream publishers. Here is the book of the article, an introduction to geology (based on *Geology* 1011X, Columbia/Barnard 1981–85) seen through the eyes of an isotope geochemist.

For Broecker, isotopes are the key to geology and he finds ways of approaching all sorts of problems from that angle. The book assumes little or no science background (it introduces powers of ten on p. 9, the visible spectrum on p. 13 and the Doppler shift on p. 17), and uses only the most elementary mathematics (the equation of a straight line on p. 105 is probably the limit). However, Broecker ingeniously manages to steer the reader from the Big Bang, through nucleosynthesis, planetary formation, planetary differentiation, moons, asteroids and comets, surface temperature and climate, to Earth resources. The final chapter is on man-environment interactions, especially the greenhouse effect.

It is, as you can see, a rather selective list, heavily biased towards the early stages of planetary evolution (it takes

nearly half of the book before the planets are constructed, let alone differentiated), but containing a good proportion of those topics that catch the imagination of students. Most results from outside geochemistry are stated, rather than argued, so that the main intellectual content is the thread of isotopic geochemistry that connects one chapter to the next. The often rather subtle isotopic arguments are presented clearly, and build one on the other through the book.

The result is something very different from other introductions to geology for the student with limited scientific knowledge, which tend to concentrate on a specific range of the usual qualitative topics. Will it sell? Will Broecker dispose of the 2,200 copies he needs to recover his \$30,000 outlay? It is a good text, well presented (but without the colour illustrations many of the big battalion of publishers seem to think necessary to attract the wilting student's attention). It is also unusual and interesting. Teachers with geochemical sympathies are likely to seek it out, rather than passively accept the latest colour-brochured offering.

In that, I think, is the answer. Conventional texts are probably best sold through conventional publishers with their excellent distribution and publicity systems for the passive purchaser. Unusual but good texts will be actively sought by people with special interests, and are the right medium for unconventional publishing. *How to Build a Habitable Planet* is a very good example of that sort of book. □

J. R. Cann is Professor of Geology in the University of Newcastle upon Tyne, Newcastle upon Tyne NE1 7RU, UK.

Physical freedom

David Bohm £27

A Primer on Determinism. By John Earman. Reidel: 1986. Pp.273. Dfl. 136, £42.50, \$64.

DETERMINISM has long been a key issue, not only in physics and philosophy, but also in a much broader context. It extends, for example, even to the question of whether human actions have any real freedom, in the sense of independence of the forces of nature. The subject has given rise to an unending controversy, to which there are as yet no signs of a clear resolution.

John Earman feels that what is most needed at this stage is a preliminary analysis, which at least brings out what is to be meant by determinism and describes the principal problems to which this notion has given rise. In doing this, he gives a survey of how determinism fares in vari-

ous branches of physics, including classical mechanics, relativity theory (special and general), probabilistic theories, modern chaos theory and the quantum theory. He draws attention to the need to distinguish carefully between the self-determination of nature and its predictability, supporting his position by a careful analysis of the general notions of lawfulness and computability, on the one hand, and chaos and randomness on the other. Eventually, he comes to the conclusion that self-determination may well be compatible with unpredictability and even with randomness, provided that the motions are unstable and possess what are, in modern chaos theory, called strange attractors.

In my opinion, the most valuable feature of the book lies in its way of pointing out how complex and delicate is the question of maintaining determinism in physical theories. Even in classical, non-relativistic physics, there are a surprising number of assumptions needed to provide for determinism, beyond the acceptance of Newton's laws. Determinism is much eas-

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

An episode from Chicago's coldest day on record, 10 January 1982, when the temperature dropped to 26°F below zero, with a severe wind chill factor, and water from hoses playing on a burning house turned to ice almost instantaneously. The picture is reproduced from *Weather* by B.W. Atkinson and Alan Gadd, a title in the new series *Earth Science Handbooks* published by Mitchell Beazley, London and Weidenfeld & Nicolson, New York. Price is £6.95, \$12.95.

ier to justify in special relativity, but to try to do this in general relativity raises a large number of further questions that have yet to be answered.

For the most part, the author accomplishes what he set out to do. However, I feel that the treatment of the quantum theory could have been improved. First, it is much too short to do justice to the subtlety of the questions that arise when one asks if the theory is compatible with determinism. Secondly, there is almost no discussion of the Copenhagen philosophy; this represents a novel approach to the whole question of determinism, which surely merited detailed consideration. In addition, there is no account of the causal interpretation (for a recent discussion see the article by B.J. Hiley and myself in *Physics Reports* 144, 323–348; 1987). The causal interpretation has been shown to be capable of providing a consistent account of self-determination (as distinguished from predictability) in the non-relativistic domain, and there are arguments indicating how such an account may be extended to a relativistic context.

Nevertheless, this book makes a useful contribution to the subject as a whole. It is well worth reading, at any rate by those who are able to understand its moderately technical and mathematical language. □

David Bohm, Department of Physics, Birkbeck College, Malet Street, London WC1E 7HX, UK, is Emeritus Professor in the University of London.

The two faces of science

Steven Yearley

Science in Action: How to Follow Scientists and Engineers Through Society. By Bruno Latour. *Open University Press/Harvard University Press: 1987. Pp.274. Hbk £25, \$25; pbk £9.95.*

IN COMMON with the recent generation of sociologists and historians of science, Bruno Latour is opposed to stories about the growth of science that explain beliefs in terms of their correspondence with nature. Rather, representations of how nature is emerge out of scientific disputes. All scientific and technical claims thus have a two-faced character: they come to look secure and impregnable where once they seemed provisional and contestable. Any particular representation of nature we come to accept is thus the consequence not the cause of agreement.

According to Latour, in making the next step from this position many historians and sociologists of science find themselves explaining representations of nature away. They replace the constraining influence of nature with that of "multi-national firms' strategies, classes, world economic trends, national cultures ... and so on" (p.256). But, of course, the same processes that reassure scientists about their beliefs operate to re-confirm for sociologists and historians the existence of cultures and classes. Latour believes "we should be as agnostic about society as about nature" (p.256).

Latour offers an alternative: as his subtitle indicates, he wants us to follow scientists and engineers through society. Starting from the scientific paper, we follow possible challenges to the authority of science into the place of scientific work, generically termed the *laboratory*. Along this route we see that there are no clear boundaries between scientific and social influences. The sources of compulsion in scientific arguments are indivisibly scientific and social. Scientists require all manner of supports and resources to make their case. Again, the decision as to which of these factors are necessary and scientific ones, and which are merely accidental, can only be made with hindsight.

Instead of forlornly seeking to sort the cognitive from the social in our accounts of science, Latour suggests that we should study how scientists and technologists enlist others to their cause. The products of science (theories, designs or whatever) gain acceptance through the establishment of networks of parties who have uses for those products. Pasteur, for example, was enmeshed in a network with "health officers, veterinary surgeons and farm

interests" (p.110). This network was mobilized not only against his scientific opponents but against diseases. Similarly, scientists working on scallops in St Brieuc Bay, Brittany, attempted to build a network with the fishing authorities, Japanese marine scientists and the scallops themselves (p.203). The things we come to regard as scientific beliefs are only as strong as the heterogeneous networks in which they are bound up.

This account of science as composed of drifting, recombining networks is presented with considerable charm and humour. There are many brief case histories to enliven the text, and the book works very well as a guide through scientific reasoning. In the end, however, I am unsure where the reader is left. Having argued that 'scientific', 'technical' and 'socio-economic' factors cannot be disaggregated, that all the influences are seamlessly linked, it is not clear what more

there is to be said or what future studies could possibly hope to demonstrate. We cannot, for example, say that such-and-such an organization of scientific labour increases the influence of economic considerations because we have to be agnostic about the status of such factors. The hostile reader may even feel that such seamlessness is equivalent to a sort of tautology.

Latour himself concludes by stating that his book offers "a breathing space to those who want to study *independently* the extensions of all [scientific and technical] networks" (p.257, my italics). Yet the main point of the book seems to be that studies and claims outside of networks are meaningless. The independent analyst is a fiction. Into whose network, then, does the book fit? □

Steven Yearley is a Lecturer in the Department of Social Studies, Queen's University, Belfast BT7 1NN, UK.

Mastering models

Jeremy Sanders

Computer-Assisted Structure Elucidation. By Neil A. B. Gray. *Wiley:1986. Pp.536. \$54.95, £52.70.*

MANY chemists enjoy the intellectual challenge of structure elucidation. To others, however, the process is merely a prelude to synthetic, mechanistic or biochemical studies, so short cuts to the structure are warmly welcomed. The most important short cuts, other than X-ray crystallography, are the spectroscopic methods: IR, NMR, UV and mass spectra all give detailed information about different aspects of molecular structure, but generally do not lead directly to a complete structure. The chemist therefore has to create a set of hypothetical structures and then test their predicted properties against the experimental evidence. The structures that fail are discarded, while those that pass are subjected to further tests. With luck, one is left with a unique correct structure. Sometimes, several candidates fit all the evidence and additional experiments are necessary; on other occasions all the structures the chemist can think of must be discarded — we tend only to propose structure types that are familiar or precedented, while nature is more creative and less prejudiced.

In this book, Neil Gray has attempted to show how computers can help in all phases of the structure-elucidation process. He explains clearly, and with abundant chemical examples, the principles behind pattern recognition, spectrum matching, and the generation and testing of structures. The book is derived from Gray's association with Djerassi's historic

DENDRAL project of the 1970s, and in places it perhaps gives undue emphasis to the details of that particular project — for example, the algorithms for generating and checking all possible structures with a particular formula are described in more detail than most chemists need. The reference lists are comprehensive, and also cover topics such as spectroscopic data bases and synthesis-analysis programs. Unfortunately the literature coverage peters out in early 1983, so there is no mention of automated analysis (using pattern recognition algorithms) of COSY NMR spectra or of NOE data.

Gray is careful to emphasize both the strengths of the computer, as in the generation of possible structures, and its weaknesses, as in the matching of spectroscopic data to unusual structural features. We all know that the computer is indispensable for collecting and processing spectroscopic data; Gray shows us that we should also use it for generating structures as candidates in the elucidation process, because that is a mechanical job the machine does better than we do. However, it is clear that, as yet, we do not know how to program the subtle interpretation of information or the inspirational leap to the right answer, presumably because we don't know how we ourselves do it.

This, then, is an interesting book which will be useful for chemists who want to know about computers or computer scientists who want to know about chemistry. It also challenges us to ponder upon, and admire, the way that the human mind goes about the business of solving problems. □

Jeremy Sanders is a Lecturer in the University Chemical Laboratory, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, UK.

The organic component in dust from comet Halley as measured by the PUMA mass spectrometer on board Vega 1

J. Kissel* & F. R. Krueger†

* Max-Planck-Institut für Kernphysik, Postfach 10 39 80, D-6900 Heidelberg 1, FRG

† Arheilger Apotheke, Messeler Str. 24, D-6100 Darmstadt 12, FRG

Following the encounter of the Vega 1 spacecraft with comet Halley, the composition of cometary dust has been analysed by mass spectroscopy. Most particles consist of a predominantly chondritic core with an organic mantle composed mainly of highly unsaturated compounds.

DUST particle impact mass spectrometers were on board the three spacecraft Vega 1, Vega 2 and Giotto, that encountered comet Halley in March 1986. Vega 1 flew through the most dusty environment, the coma, and its dust impact mass spectrometer (PUMA) was least biased by instrument effects. Therefore we concentrated on the PUMA data for the analysis of the organic components of the dust.

PUMA is a time-of-flight mass spectrometer (TOFMS) as shown in Fig. 1. A dust particle colliding with the silver target releases a large number of atomic and molecular ions from both the dust particle (projectile) and the target. Time-of-flight spectra are taken, some of them in the so-called mode-0, which means sampling the ion current every 66.67 ns at the detector. These TOF spectra are then converted into mass spectra. This technique has already been reported in detail elsewhere¹, and typical mass spectra are shown in ref. 2.

Impact ionization mass spectra

The TOFMS technique. The positive ions produced by the impact of the dust particle are accelerated and fly through a drift tube with different velocities according to their specific mass m/Z . An initial energy deviation leading to a velocity spread of ions with the same m/Z is corrected to first order by means of a reflector as described in ref. 3. As the energy distribution of the ions produced with the high impact velocity of 78 km s^{-1} (which is 34 eV per unit atomic mass specific energy) was expected to be broad, the reflector voltages were switched every 30 s such that either ions with near-zero initial energy ('long' spectra) or ions with near 150 eV energy ('short' spectra) were likely to be transmitted. 'Long' and 'short' refer to the resulting different timescales in the TOFMS per unit square root mass. By this technique at least a coarse determination of the energy distribution of the various ion types was possible.

Properties of ion formation and dust particles. Various theoretical models of ion formation postulate the creation of a hot electron-ion plasma in a Saha-equilibrium, which recombines in part during expansion. Only atomic ions have been proposed, many of them doubly charged, and even some with three or four charges. The mass spectra measured, however, exhibit atomic ions which are only singly charged, and a small fraction of molecular ions. Thus, these models cannot apply for impact ion formation. However, one of us predicted^{4,5} the ion formation process in high-velocity dust impact to be similar to other processes of rapid dissipation near solid surfaces, as, for example, sputtering used in secondary-ion mass spectrometry (SIMS). Thus we have considered collision cascade processes in SIMS as well as in dust impact. Also, in the latter, the emission of only singly charged atomic ions is expected, the number of which (Y_p from the particle, Y_t from the target) depending on the size and velocity of the incident particle and on the densities

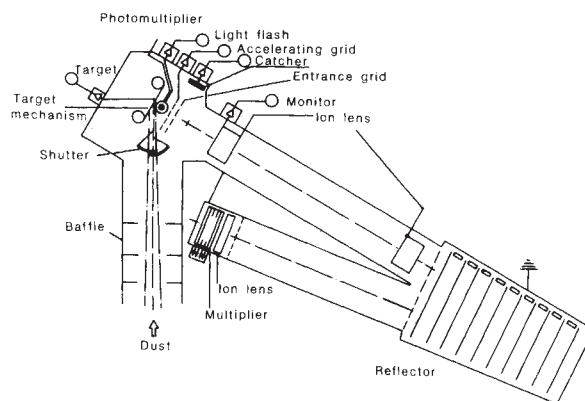


Fig. 1 The time-of-flight mass spectrometer analysing the ions released during the impact of dust particles.

of the projectile ρ_p and the target ρ_t . These dependencies have been semi-empirically deduced⁶ and are related to shock-propagation phenomena. Furthermore, Y_p also depends on the chemical nature of the projectile material, and these relative contributions are deduced from the efficiency of SIMS^{7,8} of the various elements of the periodic table. This procedure of instrumental calibration will be described in detail elsewhere⁹.

The results concerning the physical properties of the dust particles may be summarized as follows. The mass spectra and energy distributions are compatible with a typical dust particle as shown in Fig. 2. A somewhat fluffy ($\rho_p \sim 1\text{--}2 \text{ g cm}^{-3}$) mineral core, more or less chondritic, is embedded in a mantle of organic refractory material, which is even more fluffy ($\rho_p \sim 0.3\text{--}1 \text{ g cm}^{-2}$). When this dust particle impinges on the target a shock wave travels through the particle. This shock wave causes an emission of atomic and molecular species from its surface, some of which will be ionized with energies corresponding to the shock velocity. Further, the directed momentum flow of the shock wave is dissipated in the entire residual volume of the particle and the nearby target region, producing a plasma which recombines almost totally during expansion, releasing only a certain fraction of singly charged atomic ions.

In this model one can deduce not only the density but also the mass of the incident particle, if the Y_p and Y_t numbers are known. Masses between 2×10^{-15} and 10^{-11} g have been measured so far, most of them in the region between 10^{-12} g and 10^{-13} g . The absolute values of the model have been fitted to laboratory experimental points. From this procedure one can estimate the systematic error of the density determination to $\pm 20\%$, whereas that of the particle mass is within an order of magnitude.

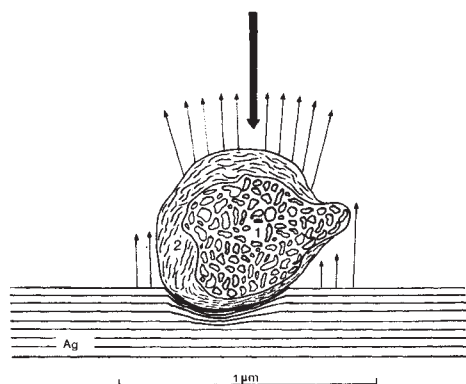


Fig. 2 Dust particle incident on the Ag target. Ion type and energy distribution shows that most dust particles consist of a fluffy silicate core (1), often covered by refractory (and icy) organic material (2), also relatively fluffy. When incident on the target with high velocity the target and dust material are compressed, and a shock wave travels through the target and particle causing the emission of atomic and molecular ions from their surfaces out of this collision cascade. Later on, the compressed material forms a plasma which recombines almost totally during expansion releasing a small fraction of singly charged atomic ions from its interior.

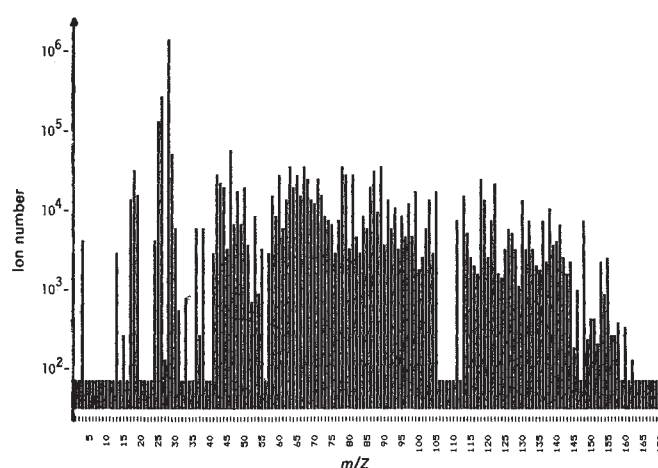


Fig. 3 Cumulative molecular ion mass spectrum on the basis of the 43 evaluated mass spectra.

Organic molecular ion formation chemistry. This report deals with the composition of the organic refractory part of the dust particles. The elemental composition of core and mantle of the dust particles as measured with PUMA-1 and evaluated with our sputtering-type model is given in Table 1. Without any bias the mean composition of the mineral core has been analysed and found to be chondritic¹⁰ within a factor of 2, on average (which was also the estimate of the systematic error of this method) and the composition of the particles as a whole was found to be practically solar (except hydrogen). However, the abundance variations of individual grains are large. Comparing with SIMS from organically contaminated surfaces, as measured in ref. 11, the ratio of the number of molecular ions to that of atomic ions in dust particle impact is reduced by at least a factor of 30. Consequently, the molecular mass spectral lines play only a minor role in the mass spectra and have to be discriminated against several types of background lines, thus affording a refined statistical analysis.

The reduction of the relative intensity of molecular ions compared with SIMS should be due to the much higher mechanical momentum flow in the impact process. Nevertheless the fact that only singly charged ions occur leads us to the conclusion that all molecular ions emitted are also electronically cold. Their formation processes are thus governed by protic reactions and a lot of skeletal rearrangements due to their high vibrational heat as produced by the shock wave.

So, one may expect organic molecular ions from the outermost layers of the dust particles to have considerable intensity only

Table 1 Mean abundances of the main elements (relative to Si = 100) in the core mantle and icy phase

Element	Abundance	Element	Abundance
Organic mantles		Chondritic cores	
H	400	C	100
C	500	O	300
N	20	Na	2
O	100	Mg	70
S	10	Al	5
		Si	100
Water ice		S	40
H	300	Ca	4
O	150	Fe	70

Errors are within a factor of two except for H, when the measurement accuracy is uncertain.

if they possess the following properties: (1) even-electron (non-radical) species (due to the non-excited electronic system); (2) stable for at least 10^{-5} s (MS flight time); (3) considerable Brønsted basicity of that related species which is expected to form a positive ion by proton addition; (4) considerable polarity (charge centres) to maintain a sufficiently low enthalpy of formation of such an ion.

These properties are known to be relevant for all such related processes of molecular ion formation⁴, such types are known to be already present in the matrix¹¹—and so this process is called 'desorption'. Bearing in mind, however, that skeletal rearrangements and fragmentations occur during desorption, we believe that we cannot determine between whether a certain ion species is a degradation product of a molecule present in the dust, or whether it is a molecular ion of a precursor molecule of reactions taking place in the solid phase of the comet itself. Nevertheless, we may be able to determine the general class of molecules forming the organic part of the dust. This is the task of the analysis procedure described here.

Analysis of molecular ions

Single and cumulative mass spectra. Those 43 mode-0 spectra of PUMA-1 which are suitable have been taken with the known atomic lines and the obvious background subtracted. Line patterns not representing a normal isotope distribution of atomic ions have been interpreted as mixtures of atomic ions of elements with normal isotope distributions and molecular ions.

At a first glance these single molecular ion mass spectra appear confusing, particularly, as one does not know the extent to which stochastic background lines contribute. Thus, further analysis is needed. The only interesting phenomenon seen in these single spectra is the 'cationization'¹³, well known with all surface ionization phenomena: silver ions from the target may react with polar molecules from the dust. These molecules cannot come from the target, because the neutral gas stream during the encounter sputters the target surface cleanly⁵ rather than causing the adsorption of molecules. Earthborne contamination is also excluded. The characteristic Ag pattern is found near ($Ag+17$), or ($Ag+18$), near ($Ag+27$) and ($Ag+30$), and thus the molecules possibly involved in this process are interpreted as being ammonia, water, hydrocyanic and formic acid, respectively.

A cumulative spectrum has been produced by adding up the linear count rates of each m/Z over all single line spectra and plotting the logarithms of these sums. The result is shown in Fig. 3. Formerly unidentified atomic ions, particularly Ti, V, Cu showing up here, have to be neglected. The widths of the K, Ca, Fe and Ag lines make these regions unidentifiable leading to the related 'holes' in the cumulative spectrum. Another

method of producing a cumulative spectrum is to sum the square roots of the count rates, thus considering each line with its statistical weight. Moreover, the frequency distribution of the occurrence of each line regardless of its intensity has also been calculated. The results are not too different, except for the ranking of the intensities. The major m/Z -lines that are attributed to molecular ions were: 18, 25, 26, 28, 29, 42, 43, 44, 60, 63, 65, 67, 71, 78, 79, 81, 87, 89, 118, 122, 130, 138 and 148. (The mass determination above the Ag peak may be unsure by 1 AMU.)

These cumulative spectra have been compared with SIMS spectra taken from targets in order to recognize possible contributions due to any remaining contamination. The lines m/Z 13(CH^+), 25(C_2H^+), 29(C_2H_3^+), 53(C_4H_3^+), 65(C_5H_3^+), 67(C_5H_7^+), 77(C_6H_5^+), 89(C_7H_5^+) and 91(C_7H_7^+) are typical contamination products, which thus cannot be attributed to dust species without further examination. (The frequent contamination ions C_2H_3^+ (27) and C_2H_7^+ (31) make the quantification of Al and P unsure; their values are thus to be taken as upper value limits only.)

Presence of lines from related ion species. The probability that two lines are present in the same mass spectrum has an almost Poisson distribution if the lines themselves are randomly distributed over the mass spectra independently from each other. On the other hand, if decomposition processes of molecules occur during ion formation, mother-daughter relations are represented as line coincidences of the related ion species, and their mass differences give the mass of the neutral elimination products. Also, the union of small molecules and the breakdown of larger molecules within the dust particle before impact cause related species to form ions simultaneously. Furthermore, reaction equilibria may be applicable between certain molecular species in the dust, the related ions of which are therefore expected to occur simultaneously in the mass spectrum. Thus, a frequency distribution of the line coincidences will exhibit such mechanisms, if one considers only those coincidences which occur much more frequently than expected from simple Poisson distribution.

In Fig. 4 this frequency distribution is plotted, each crosspoint of ordinate (lower mass no. m_1) and abscissa (mass difference no. Δm ; upper mass no. $m_2 = m_1 + \Delta m$, diagonal) marks a possible coincidence of m_1 and m_2 . If, in a statistically large number of spectra, such a coincidence occurs, this crosspoint is marked by a dot. Its size gives the frequency of the coincidences and thus its statistical significance. The smallest dot size means $n = 3$ spectra contain this coincidence (significance $s = 50\%$); the next size $n = 4$, $s = 80\%$; the next $n = 5$, $s = 92\%$; and the largest size means $n \geq 6$, $s \geq 96\%$. In the next section we report a consistent set of a few molecular species only, which can explain nearly all coincidences with $n \geq 4$.

Furthermore, we may sum all coincidences of fixed m_1 (ordinate) or m_2 (diagonal), thus resulting in the mass spectra of the ions involved in the coincidences. Contrary to the still confusing cumulative mass spectrum of Fig. 3, only a few mass lines seem to be of real chemical importance, in particular $m/Z = 18$, 26, 28, 42, 46, 48, 64, 65, 66, 68, 71, 78, 81, 89 AMU. On the other hand, we may sum all coincidences of fixed Δm (abscissa) giving a frequency distribution of neutral eliminate products. For example, $\Delta m = 2, 4, 10, 14, 17, 29, 41, 43$ AMU are clearly important, whereas $\Delta m = 18$ and 28 AMU seldom occur. Ion pairs sometimes arising at $m/Z = 17$ –19, 35–37, 53–55, 71–73, 89–91 are in part attributed to cluster ions $\text{OH} \cdot (\text{H}_2\text{O})_n^+$ and $\text{H}_3\text{O} \cdot (\text{H}_2\text{O})_n^+$, ($n = 0, 1, 2, \dots$), respectively, due to water ice residues.

Results

The ion chemistry of desorption is distinct from that of electron impact ionization in the gas phase. Whereas electron impact produces radical ion species which may decompose into radical or non-radical ions by single molecule decomposition, desorp-

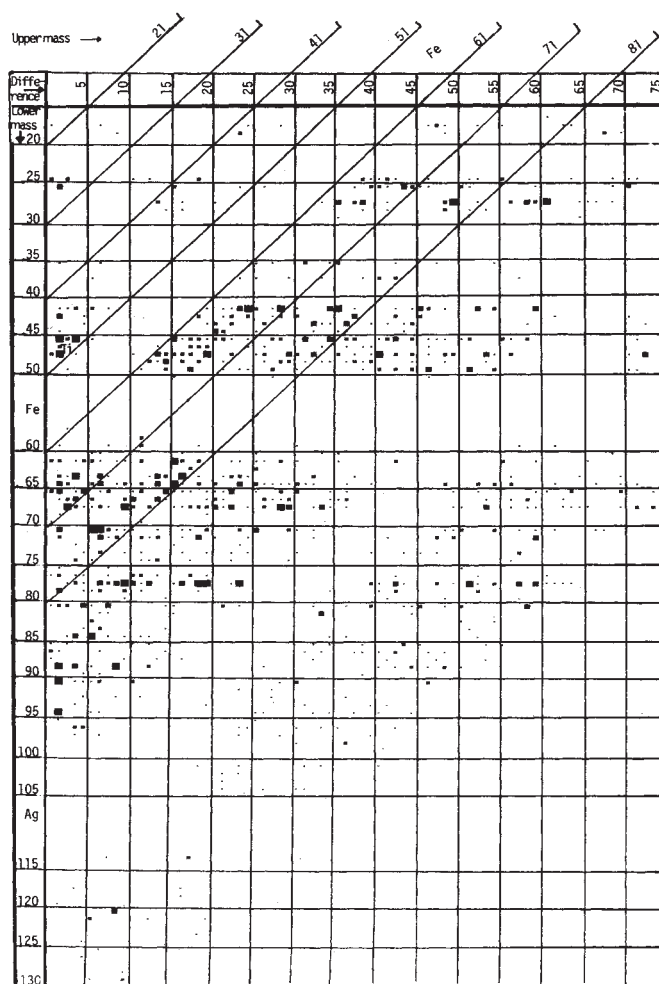


Fig. 4 Coincidence diagram of molecular mass lines. After subtracting the atomic line contributions the coincidences of each pair of mass lines occurring simultaneously on a spectrum are counted for all the spectra analysed. Frequently occurring mass pairs are marked by a dot. The larger the dot the more frequent occurrence of coincidence. Purely stochastic (Poisson) coincidences may largely contribute only to those pairs marked by the smallest dot. Thus the larger dots mark chemically significant coincidences, that is, reaction- and/or decomposition partners (as given in Fig. 5) in the dust itself or during ion formation.

tion ions are mostly non-radical ones which are pre-formed in the condensed (in this case, dust) phase through multi-molecular reaction processes. A further single molecule decomposition in the gas phase is possible; the related mechanisms obey the rules of gas-phase non-radical ions and decompose predominantly into more stable non-radical ions again. However, in TOFMS with a reflector the daughter species are observed only with lifetimes of the mother $\tau < 10^{-8}$ s, and the mother species with $\tau > 10^{-5}$ s, contrary to quadrupole- or magnet-MS. As a result, most of the desorption ions observed with TOF MS are electronically and, to a lesser extent, also vibrationally cold species.

Consequently, $m/Z = 18$ is by no means H_2O^{++} and $m/Z = 28$ is neither CO^{++} nor N_2^{++} , as would be the case with gas-phase electron impact. With desorption $m/Z = 18$ is known as the ammonia ion NH_4^+ and $m/Z = 28$ is most likely $\text{HC}\equiv\text{N}^+\text{H}$. Such ions are a frequently occurring species in the condensed phase as well as in desorption, being described as the proton-transfer result of a neutral species with Brønsted-basic properties in a polar environment. For instance, the hydrocyanic acid HCN is a weak base, exhibiting an amphoteric character and thus forming H_2CN^+ ; however, it is a stronger acid, forming most likely the carbanion CN^- which we do not observe because it has a negative charge. It is less likely, due to its electronically

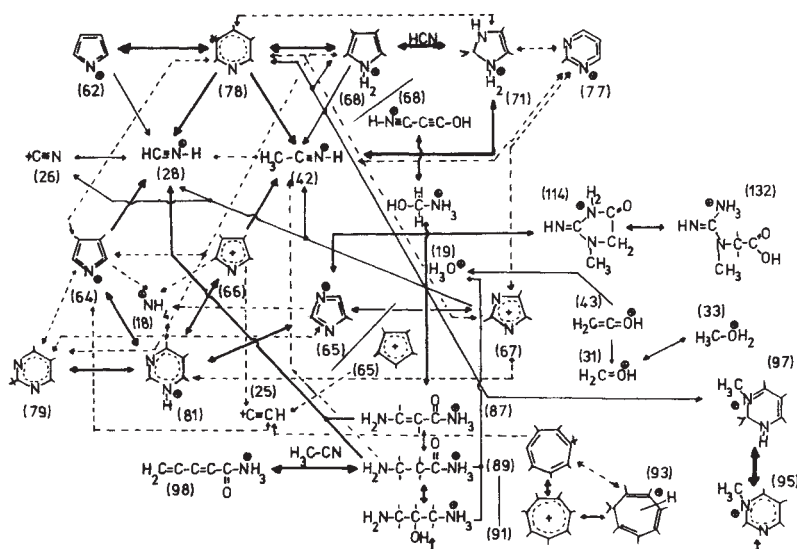


Fig. 5 Reactions and decomposition processes of various ions as consistent with the frequently occurring mass line coincidences as depicted in Fig. 4. Also consistent with organic molecular ion formation processes known from other techniques with fast dissipation of energy near solid surfaces.

unsaturated property, to form the carbenion $\text{N}\equiv\text{C}^+$, but this was found to a lesser extent. The presence of the ions related to ammonia and hydrocyanic acid does not necessarily mean that these substances are present in the dust, however, it does point to the presence of nitrogen-containing hydrocarbons. The first one especially indicates the presence of amines, the second one nitriles, imines and heterocyclic aromates. Water or hydrates would form H_3O^+ ($m/Z=19$), and carbohydrates COH^+ ($m/Z=29$) or COH_3^+ ($m/Z=31$), which are found to a much lower extent. Carbohydrides form preferably C_2H_3^+ ($m/Z=25$), especially if unsaturated, otherwise they form C_2H_3^+ ($m/Z=27$, undistinguishable from Al), but only with low probability due to its low polarity.

A very important mass number in the spectra, $m/Z=42$, is attributed to the acetonitril ion $\text{H}_3\text{C}-\text{C}\equiv\text{N}^+\text{H}$. Like the hydrocyanic cation $m/Z=28$, this ion does not only indicate the presence of acetonitrile but also the same class of substances mentioned above.

With $m/Z=44$ the imino ion $\text{H}_3\text{C}-\text{CH}=\text{N}^+\text{H}_2$ as well as the H_2CNO^+ (due to the low basicity of the (iso-) cyanic acid HCNO or HOCN , respectively) may both contribute, and with $m/Z=46, 48$ further reduction products of both ions may be present. (However, coincidences between $m/Z=46, 48, 50$ indicate the presence of titanium atomic ions which were not considered before analysis.) The organic molecular part of $m/Z=48$ can only be explained as the aminomethanol ion $\text{HO}-\text{CH}_2-\text{NH}_3^+$ by coincidence analysis, as it tends to disintegrate under water elimination to the iminomethane ion $\text{H}_2\text{C}=\text{N}^+\text{H}_2$ ($m/Z=30$). In general, their lower intensities indicate that they are only a small proportion of the oxygen-containing ion types.

The same is true when looking at the neutral eliminated products. An elimination of water from a molecular ion m_2 leading to m_1 with $m_2 - m_1 = \Delta m = 18$ is not found to contribute significantly. Also the elimination of CO ($\Delta m = 28$), well known for ketones, esters, furanes, or chinones does not contribute. On the contrary, the eliminations of NH_3 ($\Delta m = 17$, ammonia), $\text{H}_2\text{C}=\text{NH}$ ($\Delta m = 29$, methylimine), and that of $\Delta m = 41$ (acetonitrile $\text{H}_2\text{C}-\text{C}\equiv\text{N}$), and $\Delta m = 43$ (either $\text{H}_3\text{C}-\text{CH}=\text{NH}$ (ethanimine) or tautomerically $\text{H}_2\text{C}=\text{C}-\text{NH}_2$ (ethenamine)) are found to be frequent reactions. This all points to the presence of unsaturated N-containing carbohydrides.

The different oxidation-reduction states of these unsaturated compounds correspond to the common values of $\Delta m = 2$ and 4, interpreted as being elimination (oxidation) processes of one or two H_2 molecules, forming -enes, dienes, or -ines. So we are able to present a consistent interpretation of the main low mass lines, coincidences and mass differences. The presence of unsat-

urated nitrogen-containing species like nitriles, aldimines, enamines and possibly their polymerization products is assured. The latter ones are the main types of species which are able to explain all the major coincidences as shown in Fig. 5—pyrrol, pyrazol/imidazol, pyridine, pyrimidine and its derivatives are very likely to be present in the dust.

Due to the low fraction of nitrogen relative to carbon found in the elemental composition, it may be surprising at first sight not to find the unsaturated carbohydrides to a larger extent. However, their probability of forming ions in desorption is a factor of ~ 100 lower than the related nitrogen-containing ones, because they lack a charge centre and thus have only a low basicity. Furthermore, they are consequently more volatile, so they may be evaporated out of the dust before impact. Nevertheless, we conclude that a considerable fraction of alkenes, alkynes, or cyclic, especially aromatic, molecules (like benzene, toluene) may well be present, as the coincidences of $m/Z=89, 91, 93$ indicate, which may be correlated with the most stable tropylium-ion (91) and its oxidation product dehydrotropylium as well as the reduction product the cycloheptatriene (93) ion.

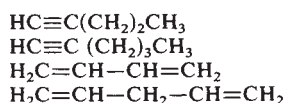
With larger m/Z the interpretation becomes ambiguous. The presence of oxygen atoms in the molecules may well occur more frequently with larger molecules, but this has not been proved. With a lack of oxygen there are even some indications from multiple coincidences in a few spectra that a molecule as large as adenine may be present. Its quasi-molecular ion ($m/Z=136$) is completely oxygen-free, but as an aminopurine it belongs to the class of heterocyclic aromates discussed above.

It is therefore consistent that no indications for alcohol or sugars as N-free carbohydrates have been found, nor α -amino acids ($\text{R}-\text{CH}-\text{NH}_2-\text{COOH}$), although the latter ones are known to produce large yields of ions, easily detectable due to their common elimination of formic acid ($m=46$). So the biologically important α -amino acids are, if they are formed at all, at least a factor of 30 less abundant than the biologically important pyrimidines and purines.

It is puzzling that the presence of oxygen in the organic molecules was not found to a larger extent, although there is much oxygen within the dust, as seen from the elemental composition. Possibly, most of the oxygen is bound in the silicates, and also present as frozen H_2O and CO_2 . However, we found indications that (Ag-cationized) formic acid was present, and possibly also some other N-free carboxylic acids as acetic or oxalic acid may well be formed. Also methanol and/or ethanal may be present. Their efficiencies of producing positive ions are fairly low. With a strong acidic behaviour they are expected to show up in the negative ion spectrum, which we did not measure. The same is true for S-containing species, which tend to form

Table 2 Types of organic molecules as inferred from the mass spectra**C—H— Compounds**

(Only high-molecular probable due to volatility; hints only to unsaturated)



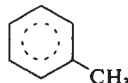
Pentyne
Hexyne
Butadiene
Pentadiene



Cyclopentene, cyclopentadiene



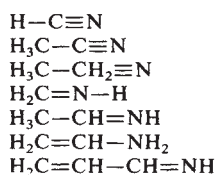
Cyclohexene, cyclohexadiene



Benzene, toluene

C—N—H— Compounds

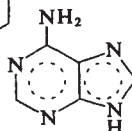
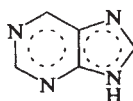
(Mostly of high extensity; also higher homologues possible)



Hydrocyanic acid
Ethanenitrile (acetonitrile)
Propanenitrile
Iminomethane
Iminoethane
Aminoethene } (tautomeric)
Iminopropene



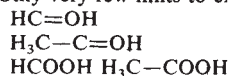
Pyrroline, pyrrole, imidazole



Pyridine, pyrimidine
(and derivatives)
Purine, adenine

C—O—H— Compounds

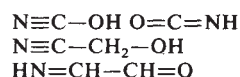
(Only very few hints to existence)



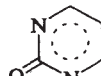
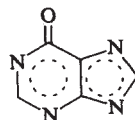
Methanal (formaldehyde)
Ethanal (acetaldehyde)
Methanoic (formic) and
ethanoic (acetic) acid (?)

C—N—O—H— Compounds

(Amino-, Imino-, Nitrile of -ole, -ale, -keto- only probable with higher C-numbers of -anes, -enes, and -ines or cyclic aromates)



(Iso-) cyanic acid (?)
Methanol nitrile
Methanalimine } (tautomeric)



Oxyimidazole, oxypyrimidine

Xanthine

Structure isomers are additionally possible. Several types may form tautomers, mesomers and conformational isomers. Thus the molecules given here serve only as examples of the class of substances possibly present in the organic component of the dust. We are not sure yet if oxygen-containing species are present.

more negative ions from thiols, thials and so on. The lack of coincidences for $\Delta m = 34$ yielding a possible H_2S elimination analogous to that of water do not point to the presence of such species. The substance classes that are probably present in comet Halley's dust are summarized in Table 2.

Possible implications

The hypothesis that the organic part of the dust of a comet consists of sperrmae¹⁴ (F. Hoyle and N. C. Wickramasinghe,

preprint no. 122, University College, Cardiff, 1986) has been disproved at least for that of comet Halley as analysed by Vega 1. First of all, dried cells contain at least some fraction of Na and K: we found only less than 1 per 10^6 in the organic dust particles. No indications have been found for a significant presence of oxygen-rich compounds such as sugars and peptides, although the latter molecules should show up in the mass spectra if present. The value $m/Z = 31$ is found to be of minor importance—it is due to phosphorus in the mineral core only. However,

nucleic acids should yield phosphorus, but we do not see it in the organic fraction sufficiently; also sugars are known to produce $m/Z = 31$ as CH_2OH^+ ion.

Nevertheless we do think that the question of the origin of life in the context of primordial matter as found in comets has become even more exciting. We stress that the substance classes, which we believe are present in the cometary dust are highly reactive especially in warm water. Imagine that some mechanism brings such substances into contact with liquid water. The unsaturated carbohydrides may add water molecules thus reacting to give carbohydrates such as sugars. The nitrogen-containing species may react to give bases of nucleosides, if not already present. The mineral core may serve as the phase boundary necessary to give a local concentration gradient, and also as a

source for dissolved phosphoric acid. With such a model, the chemical prerequisites and also the requirements of nonlinear non-equilibrium thermodynamics are fulfilled to trigger, for example, hypercyclic behaviour¹⁵. High affinity, high concentration gradient and localization of the reactions have been considered^{16,17} to trigger the self-organization of nucleic acids. This is known to be possible without amino acids at this early stage¹⁵, because the replication of RNA is not necessarily associated with transcription¹⁸, when Zn^{2+} -catalysts are present, as is the case with the dust.

We thank Mrs R. Krueger for critical comments and laboratory support in organic chemistry for mass spectrometric model investigations. The work has been supported in part by BMFT and ESA.

Received 26 January; accepted 5 March 1987.

1. Kissel, J. *Eur. Space Agency Spec. Publ.* **1077**, 67 (1986).
2. Kissel, J. *et al. Nature* **321**, 280 (1986).
3. Mamyrin, B. A., Karatyev, V. I., Shmikk, D. V. & Zagulin, V. A. *Zh. éksp. teor. Fiz.* **64**, 82 (1973).
4. Krueger, F. R. *Z. Naturf.* **38a**, 385 (1983).
5. Krueger, F. R. *Eur. Space Agency Spec. Publ.* **224**, 49 (1984).
6. Krueger, F. R. *Eur. Space Agency Spec. Publ.* **224**, 43 (1984).
7. Sparrow, G. R. *Am. Soc. for Mass Spectrom.* **24** (San Diego, 1976).

8. Kissel, J. & Krueger, F. R. *Appl. Phys.* **A42**, 69 (1987).
9. Kissel, J. *et al.* (in preparation).
10. Jessberger, E. K., Kissel, J., Fechtig, H. & Krueger, F. R. *Eur. Space Agency Spec. Publ.* **249**, 27 (1986).
11. Benninghoven, A. & Sichtermann, W. K. *Analyt. Chem.* **50**, 1180 (1978).
12. Renner, D. & Spittler, G. IFOS III, Springer Ser. in Phys. **9**, 126 (1986).
13. Schueler, B., Feigl, P., Krueger, F. R. & Hillenkamp, F. *Org. Mass Spectrom.* **16**, 502 (1981).
14. Eigen, M. & Schuster, P. *The Hypercycle* (Springer, Berlin, 1977).
15. Nicolis, G. & Prigogine, I. *Self-Organization in Non-equilibrium Systems* (Wiley, New York, 1977).
16. Krueger, F. R. *Physik und Evolution* (Parey, Berlin, 1984).

Acid-dependent ligand dissociation and recycling of LDL receptor mediated by growth factor homology region

C. Geoffrey Davis, Joseph L. Goldstein, Thomas C. Südhof, Richard G. W. Anderson, David W. Russell & Michael S. Brown

Departments of Molecular Genetics, Internal Medicine and Cell Biology, University of Texas Health Science Center at Dallas, Southwestern Medical School, 5323 Harry Hines Boulevard, Dallas, Texas 75235, USA

A domain in the low-density lipoprotein receptor contains three cysteine-rich 'growth factor' repeats like those that occur in many proteins. When this domain is deleted, the receptor no longer releases its ligand at acid pH, it is no longer recycled efficiently and it is rapidly degraded after ligand binding.

A SURPRISING variety of extracellular proteins contain multiple copies of a 40-amino acid, cysteine-rich sequence that is derived from a common ancestor. This sequence was first noted in the precursor for epidermal growth factor (EGF), a peptide hormone¹. Similar 'growth factor' repeats have now been identified in cell-surface receptors (the low density lipoprotein (LDL) receptor and thrombomodulin); viral proteins (vaccinia virus P19); circulating proteins of the blood clotting system (factors IX and X, proteins C and S, urokinase and tissue plasminogen activator); growth factors (EGF and tumour growth factor- α); and developmental proteins of lower eukaryotes (products of the *notch* locus of *Drosophila* and the *lin-12* locus of *Caenorhabditis elegans*)¹⁻⁷. These proteins contain up to 36 growth factor repeats, each of which contains six cysteine residues that are thought to form three internal disulphide bonds. The only obvious feature that these proteins hold in common is their location in the extracellular environment, either as membrane-bound or as secreted glycoproteins. The function of the growth factor repeats is not known at present for any of the above proteins.

The most compelling evidence for a common evolutionary origin of the growth factor repeats comes from a comparison of the genes for the human LDL receptor and the human EGF precursor². One half of the amino acids in the human LDL receptor are homologous to a 400-amino-acid region in the human EGF precursor. This shared sequence is encoded by

eight exons; five of the exon-intron boundaries are identical in the two genes, implying that these exons form a duplicated cassette that was transferred *en bloc* between genes². Three of the shared exons encode growth factor repeats; two of them occur in tandem at the amino-terminus of the 400-amino-acid segment and one occurs at the carboxy-terminus (Fig. 1).

The LDL receptor may provide a good model system by which to discover the function of growth factor repeats because this protein has a modular structure that allows various domains to be deleted without abolishing all functions of the receptor^{8,9}. This protein belongs to a class of recycling receptors that bind ligands at the surface, internalize them in coated pits, dissociate from the ligands in endosomes and are returned to the cell surface¹⁰. The LDL receptor, like many other recycling receptors, separates from its ligand at acid pH. Intracellular dissociation is mediated by proton pumps within the wall of the endosome that lower the pH (refs 11 and 12).

In these studies we have used a deletion analysis to show that one function of the EGF precursor region in the LDL receptor is to accelerate acid-dependent ligand dissociation and thereby to facilitate receptor recycling.

Expression of deleted LDL receptor

Figure 1 shows the five domains of the LDL receptor, extending from the extracellular amino-terminal alanine to the intracellular carboxy-terminal alanine (residue 839). The first domain con-

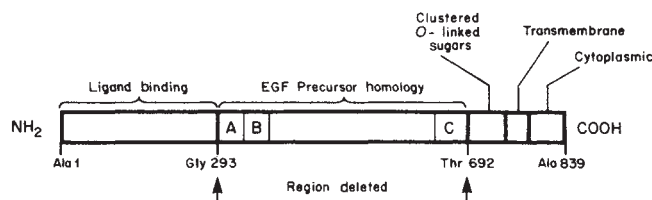


Fig. 1 The domain structure of the human LDL receptor is shown. The three growth factor repeats (A, B and C) in the EGF precursor homology region are indicated. Plasmid pLDLR2 contains a cDNA corresponding to the entire coding region under control of the SV40 early region promoter. Plasmid pΔEGFP was derived from pLDLR2 by deletion of the DNA encoding amino acids glycine 293 to threonine 692, which correspond to the entire EGF precursor homology region⁸. The borders of the deletion correspond precisely to exon-intron boundaries in the gene.

Methods. For plasmid construction¹⁴, a mutagenic oligonucleotide (42 nucleotides long) that bridged between nucleotide positions 920 to 939 and nucleotides 2,140 to 2,161 in pLDLR2 (ref. 13) was annealed to a single-stranded DNA template isolated from an M13mp18 recombinant clone containing a portion of pLDLR2 (ref. 9). The cDNA insert in the M13 clone extended from the *EcoRI* site at amino acid 219 (nucleotide 750) in the ligand binding domain to the *SstI* site (nucleotide 4,086) in the 3' untranslated region¹³. After primer extension and ligation, aliquots of the mutagenesis reaction were transformed into *E. coli* TG-1 cells, and the resulting plaques were screened by hybridization with the ³²P-labelled mutagenic oligonucleotide. After confirmation of the mutation by DNA sequencing, single-stranded M13 DNA was converted to double-stranded DNA by primer extension²⁴. The double-stranded cDNA insert containing the deletion was excised by digestion with *EcoRI* and *SstI*. The insert was purified by polyacrylamide gel electrophoresis and reinserted into the same sites of pLDLR2. The structure of the resulting plasmid was verified by restriction enzyme mapping and DNA sequencing²⁴. For isolation of transfected cell lines, stable cell lines were derived from antibiotic G418-resistant colonies after cotransfection⁹ of plasmids containing LDL receptor cDNAs together with pSV3-Neo into *ldla*-7 cells, a mutant Chinese hamster ovary cell line lacking functional LDL receptors¹⁵. Two stable cell lines were obtained: TR715-19 cells, which were transfected with pLDLR2, and TR841a-2 cells, which contained pΔEGFP. Cell lines were maintained at 37 °C in 5–7% CO₂ in monolayers in medium A (Ham's F-12 medium containing 2 mM glutamine, 100 U ml⁻¹ penicillin, 100 μg ml⁻¹ streptomycin) supplemented with 4% (v/v) newborn calf lipoprotein-deficient serum, 1% (v/v) fetal calf serum (FCS), 20 μM compactin, and 0.2 mM sodium mevalonate⁹. For immunoprecipitation of ³⁵S-labelled LDL receptors and assays of LDL receptor activity (Figs 2 and 4–8), 4–6 × 10⁴ cells from trypsinized stock cultures were seeded into 60-mm Petri dishes in medium A supplemented with 10% FCS (day 0). On day 2, cells were re-fed with medium of identical composition. On the day before experiments (day 3 or 4), cells received medium A supplemented with 10% calf lipoprotein-deficient serum, 0.1 μg ml⁻¹ 25-hydroxycholesterol, and 10 μg ml⁻¹ cholesterol to suppress any endogenous defective hamster LDL receptors. Experiments were performed on day 4 or 5. For immunoblotting experiments (Fig. 3), 6–7 × 10⁴ cells were seeded into 100-mm Petri dishes on day 0, and the cells were fed with the same medium as above.

tains the ligand binding site. The second domain contains the region with homology to the EGF precursor. The third domain is rich in serine and threonine residues and contains clustered O-linked carbohydrate chains. The fourth domain is the hydrophobic transmembrane region. The fifth domain is the cytoplasmic tail that directs the receptor to coated pits¹⁰. The current experiments were performed with a full-length complementary DNA for the LDL receptor in an expression vector driven by the Simian virus 40 (SV40) early region promoter¹³. Through oligonucleotide-directed mutagenesis¹⁴, we deleted the DNA sequences that encode the 400 amino acids that comprise the EGF precursor homology region (glycine 293 to threonine 692). The plasmid bearing the deletion was introduced by transfection

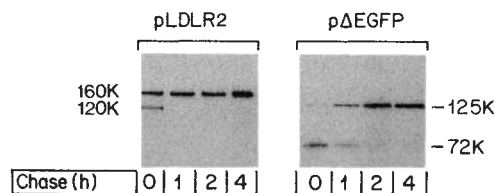


Fig. 2 Electrophoresis of ³⁵S-labelled LDL receptors immunoprecipitated from transfected cells. On day 5 of cell growth, two dishes of cells expressing the indicated cDNA were incubated in methionine-free medium⁹ containing 200 μCi ml⁻¹ ³⁵S-methionine for 2 h (pulse). The medium was then switched to complete medium (0.2 mM methionine) for the indicated time (chase). Radiolabelled LDL receptors were precipitated from detergent extracts by incubation with immune complexes containing the monoclonal antibody IgG-C7 directed against the normal receptor¹⁶. Immunoprecipitates were processed for electrophoresis on SDS/7% polyacrylamide gels¹⁶, and the dried gels exposed to XAR-5 film for 5 days (pLDLR2) or 7 days (pΔEGFP) at -70 °C. Apparent sizes (*M_r*) of the radiolabelled proteins were calculated from positions of the following standards as determined by Coomassie blue staining: myosin (200K), β-galactosidase (116K), phosphorylase B (97K), bovine serum albumin (68K) and ovalbumin (45K)¹⁶.

into *ldla*-7 cells, a line of mutagenized Chinese hamster ovary cells that do not produce functional LDL receptors¹⁵. Cells expressing high levels of the transfected LDL receptor cDNA were selected through immunofluorescent labelling with an antibody that is specific for the human LDL receptor⁷.

Permanent lines of cells transfected with either the normal receptor cDNA (pLDLR-2) or the deleted cDNA (pΔEGFP) were pulse-labelled for 2 h with ³⁵S-methionine and then switched to unlabelled methionine for various times (Fig. 2). The LDL receptor was immunoprecipitated with IgG-C7, a monoclonal antibody that recognizes the ligand binding domain¹⁶. After electrophoresis in sodium dodecyl sulphate (SDS) gels, the precursor form of the normal receptor was visualized as a band of relative molecular mass 120,000 (*M_r*, 120K). It was rapidly processed to a mature form that migrated at 160K, an anomalous increase that is attributable to extension of the O-linked carbohydrates in the Golgi complex⁹. The precursor produced by pΔEGFP migrated at 72K, and it was rapidly converted to a mature form of 125K. Thus, despite the extensive deletion, the mutant LDL receptor was apparently transported at a nearly normal rate to the Golgi complex, where the O-linked sugars underwent normal processing.

Ligand binding

Extracts of transfected cells were subjected to SDS gel electrophoresis, transferred to nitrocellulose, incubated with various ligands and visualized by autoradiography or enzymatic colour development (Fig. 3). Electrophoresis was performed in the absence of sulphhydryl reducing agents, and therefore the apparent *M_r*s of the mature receptors were lower than those in Fig. 2 (ref. 17). The normal receptor and the mutant receptor bound monoclonal antibody IgG-C7 as expected. Both receptors also bound two ligands, β-VLDL and LDL, which bind by virtue of two different apoproteins^{18,19}. The active moiety of LDL is apoB-100, a large protein of ~514K that occurs in one copy per lipoprotein particle. The active species in β-VLDL is apoE, a small protein of 33K that occurs in multiple copies per particle. Both proteins bind to the ligand binding domain, and each competes with the other. The normal LDL receptor has a higher affinity for apoE than for apoB-100¹⁹.

Although the deleted receptor bound both ligands on nitrocellulose blots (Fig. 3), in intact cells it showed high-affinity binding of β-VLDL (Fig. 4b) but not LDL (Fig. 4a). To estimate the stoichiometry of binding, we normalized the data for the binding of ¹²⁵I-labelled IgG-C7, which we took as a measure of the total number of receptors on the cell surface. When this correction

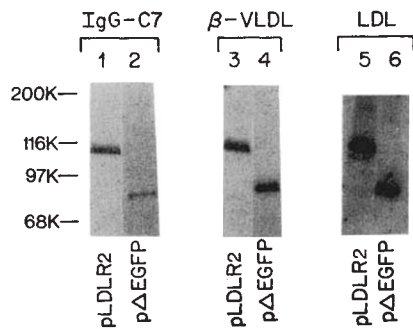


Fig. 3 Immunoblotting (lanes 1 and 2) and ligand blotting (lanes 3–6) of solubilized LDL receptors from transfected cells. On day 5 of cell growth, cells expressing the indicated cDNA from one 100-mm dish were harvested, and an aliquot of the detergent-solubilized extract (350 μ g protein) was subjected to electrophoresis on SDS/7% polyacrylamide gels under nonreducing conditions and transferred to nitrocellulose filters¹⁷. Lanes 1, 2, the filter was incubated for 2 h with 10 μ g ml⁻¹ of mouse monoclonal antibody IgG-C7 against normal receptor, washed²⁵, and then incubated with goat anti-mouse immunoglobulin G (IgG) coupled to alkaline phosphatase and colour development reagents (Bio-Rad Immun-Blot assay kit). Lanes 3, 4, the filter was incubated for 1 h with 5 μ g protein per ml of ¹²⁵I-labelled β -VLDL ($\sim 5 \times 10^5$ c.p.m. per μ g protein), washed, and exposed to X-ray film at -70°C for 1 h with an intensifying screen²⁵. Lanes 5, 6, the filter was incubated for 1 h with 7.5 μ g protein per ml of human LDL, followed by incubation for 1 h with 5 μ g ml⁻¹ of rabbit anti-human apoB (IgG fraction, Calbiochem; purified by protein A-Sepharose chromatography), followed by incubation for 30 min with ¹²⁵I-labelled goat anti-rabbit IgG ($\sim 5 \times 10^6$ c.p.m. ml⁻¹; $\sim 3.5 \times 10^6$ c.p.m. μ g⁻¹)¹⁷. The filter was washed and exposed to X-ray film at room temperature for 18 h. Size standards are M_r as before (see Fig. 2).

was made, the deleted receptor bound a normal amount of β -VLDL (Fig. 4d), but a markedly reduced amount of LDL (Fig. 4c).

Internalization of ¹²⁵I-labelled β -VLDL

Cells transfected with the normal and p Δ EGFP cDNAs were incubated with ¹²⁵I-labelled β -VLDL at 4°C to obtain cell surface binding (Fig. 5). The cells were washed, warmed to 37°C for various intervals, then chilled to 4°C , and the amount of surface-bound β -VLDL was measured by its susceptibility to release with a polyanion, suramin²⁰. We also measured the amount of radioactivity that was resistant to suramin release (internalized β -VLDL), and we measured the amount that had been degraded to trichloroacetic acid-soluble material. The kinetics of internalization and degradation were indistinguishable in cells transfected with the two cDNAs. In both cases approximately half the surface-bound material entered the cell within 20 min. After a 30-min lag, rapid degradation was observed (Fig. 5).

Receptor recycling

Cells transfected with normal and p Δ EGFP cDNAs were incubated with ¹²⁵I-labelled β -VLDL continuously at 37°C , after which the cells were rapidly chilled to 4°C , washed, and the amount of surface-bound (suramin-releasable), internalized (suramin-resistant), and degraded ¹²⁵I-labelled β -VLDL was measured (Fig. 6). A striking difference in the behaviour of the two cell lines was noted. In cells that expressed the normal cDNA the amount of surface-bound ¹²⁵I-labelled β -VLDL remained at a relatively constant steady state level (solid circles, Fig. 6a). The amount of internalized ¹²⁵I-labelled β -VLDL rose rapidly during the first hour and then more slowly thereafter (Fig. 6b). Degradation continued linearly over the 6 h, indicating that the cells were continuing to bind and internalize the β -

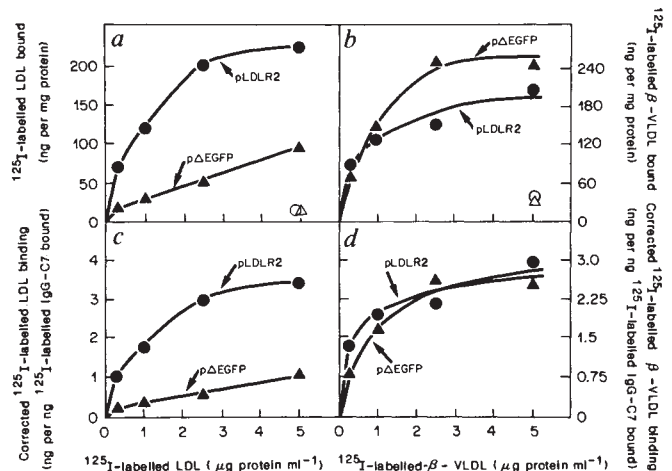


Fig. 4 Surface binding of ¹²⁵I-labelled LDL (a and c) and ¹²⁵I-labelled β -VLDL (b and d) to monolayers of cells transfected with pDLR2 (●) or p Δ EGFP (▲) at 4°C . On day 3 of cell growth, cells expressing the indicated cDNA received 1.5 ml of ice-cold medium B (Eagle's minimum essential medium (MEM) without bicarbonate, (Gibco) containing 5% human lipoprotein-deficient serum and 20 mM HEPES at pH 7.4) and the indicated concentration of either ¹²⁵I-labelled LDL (385 c.p.m. per ng protein) (a) or ¹²⁵I-labelled β -VLDL (449 c.p.m. per ng protein) (c) in the absence (●, ▲) or presence (○, △) of 350 μ g protein per ml of the corresponding unlabelled lipoprotein. After incubation for 3 h at 4°C , the cells were washed rapidly six times with ice-cold buffer C (50 mM Tris-Cl, 165 mM NaCl, 3 mM CaCl₂, and 0.2% (w/v) bovine serum albumin at pH 7.4), and the total radioactivity bound to the cells was determined as described²⁶. In the same experiment, the cell surface binding of ¹²⁵I-labelled IgG-C7 was measured by incubating separate cell monolayers with 2 μ g ml⁻¹ of ¹²⁵I-labelled IgG-C7 (793 c.p.m. ng⁻¹)²⁷. c and d, Corrected values calculated by dividing the amount of bound ¹²⁵I-labelled LDL or ¹²⁵I-labelled β -VLDL by the amount of bound ¹²⁵I-labelled IgG-C7. Cells transfected with pDLR2 and p Δ EGFP bound 70 and 95 ng of ¹²⁵I-labelled IgG-C7 per mg cell protein, respectively (mean of triplicate incubations). Each value for ¹²⁵I-labelled LDL and ¹²⁵I-labelled β -VLDL is a single incubation except for values at 5 μ g protein ml⁻¹, which are the average of duplicate incubations. Human LDL (density, 1.019–1.063 g ml⁻¹) and rabbit β -VLDL (density, <1.006 g ml⁻¹) were prepared by ultracentrifugation and radiolabelled with ¹²⁵I by the iodine monochloride method²⁶. ¹²⁵I-labelled IgG-C7 was prepared and radiolabelled with ¹²⁵I as described²⁷.

VLDL at a steady rate (Fig. 6c). In contrast, in the p Δ EGFP-transfected cells the amount of binding reached a peak at 10 min and then declined over the next 3 h (open circles in Fig. 6a). The amount of internalized β -VLDL reached a maximum at about 1 h and then failed to increase (Fig. 6b). The amount of degradation reached a plateau after 3 h, reflecting the reduction in cell surface β -VLDL binding sites (Fig. 6c). These data suggested that in the presence of β -VLDL the receptor encoded by p Δ EGFP was gradually lost from the cell surface.

To study the loss of receptors by a different technique, we incubated transfected cells at 37°C with unlabelled β -VLDL or LDL (Fig. 7, upper panels). After various times, the cells were chilled, washed, treated with suramin to remove bound β -VLDL, and then incubated at 4°C with ¹²⁵I-labelled IgG-C7 to measure the number of surface LDL receptors. Unlabelled LDL produced little change in the number of cell surface receptors over a 5-h period in cells expressing either the normal cDNA or p Δ EGFP (Fig. 7b). β -VLDL produced a profound decline in LDL receptors in the p Δ EGFP-transfected cells (Fig. 7a). Cells transfected with the normal cDNA also showed some decline in receptors on incubation with β -VLDL, but the decline was much less than that seen in the cells transfected with p Δ EGFP (Fig. 7a).

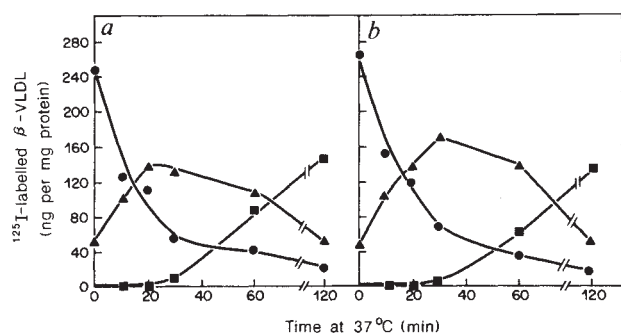


Fig. 5 Internalization and degradation at 37 °C of ^{125}I -labelled β -VLDL previously bound to LDL receptors at 4 °C in monolayers of cells transfected with *a*, pLDLR2 or *b*, p Δ EGFP. The time course follows labelled β -VLDL that was surface bound (●), internalized (▲) and degraded (■). On day 4 of cell growth, triplicate monolayers of cells expressing the indicated cDNA received 1.5 ml of ice-cold medium B supplemented with 5 μg protein per ml of ^{125}I -labelled β -VLDL (232 c.p.m. per ng protein) in the absence or presence of 370 μg protein per ml of unlabelled β -VLDL. After 2 h at 4 °C, each monolayer was rapidly washed six times with ice-cold buffer C. Each dish then received 2 ml of warm medium D containing 5 μg protein per ml of unlabelled β -VLDL, and all dishes were incubated at 37 °C. After the indicated time, groups of dishes were rapidly chilled to 4 °C, the medium was removed, and its content of ^{125}I -labelled trichloroacetic acid-soluble material (■) was measured. Each cell monolayer was washed rapidly six times with ice-cold buffer C, after which the amounts of surface-bound ^{125}I -labelled β -VLDL (suramin-releasable fraction) (●) and internalized ^{125}I -labelled β -VLDL (suramin-resistant fraction) (▲) were determined by modification of a previously described release assay²⁶ in which suramin²⁰ was used as the anionic releasing agent in place of dextran sulphate. Each monolayer was incubated for 1 h at 4 °C with 2 ml of buffer containing 10 mM suramin, 50 mM NaCl, 3 mM CaCl_2 , and 10 mM HEPES at pH 7.4, after which surface-bound and internalized ^{125}I -labelled β -VLDL were separately measured²⁶. The values shown represent high affinity values, which were determined by subtracting the values for ^{125}I -labelled β -VLDL in the presence of unlabelled β -VLDL (nonspecific value) from that in its absence (total value).

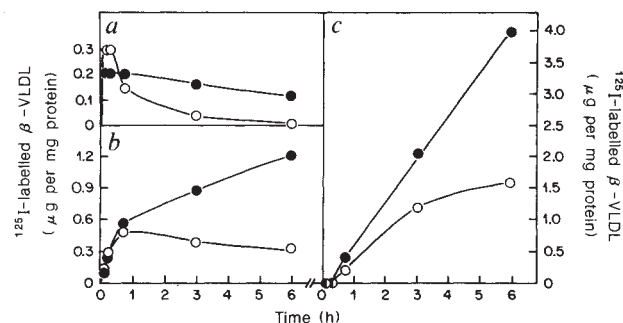


Fig. 6 Time course of metabolism of ^{125}I -labelled β -VLDL in cells transfected with pLDLR2 (●) or p Δ EGFP (○), showing β -VLDL that was *a*, surface-bound; *b*, internalized or *c*, degraded. On day 5 of cell growth, duplicate monolayers of cells expressing the indicated cDNA received 2 ml of warm medium D (Eagle's MEM with NaHCO_3 and without glutamine (Gibco) containing 5% human lipoprotein-deficient serum) and 10 μg protein per ml of ^{125}I -labelled β -VLDL (256 c.p.m. per ng protein) in the absence or presence of 500 μg protein per ml of unlabelled β -VLDL. After incubation at 37 °C for the indicated time, high affinity values were determined for the amount of ^{125}I -labelled β -VLDL that was surface bound, internalized, and degraded by the cells. Surface-bound and internalized ^{125}I -labelled β -VLDL were measured by the suramin release assay (see Fig. 5); degraded ^{125}I -labelled β -VLDL was measured by trichloroacetic acid precipitation²⁶.

increase in ligand release at both 37 °C and 4 °C. In contrast, there was no acid-dependent release from the receptor encoded by p Δ EGFP at either temperature. In the same experiment, suramin (10 mM) at pH 7.4 at 4 °C released 86 and 90% of the receptor-bound ^{125}I -labelled β -VLDL in the pLDLR2- and p Δ EGFP-transfected cells, respectively. The ligand was also released from the mutant receptor by EDTA, indicating that the EGF precursor region is not responsible for the calcium dependence of binding (data not shown).

Functions of growth factor repeats

Our deletion analysis demonstrates that the EGF precursor homology region in the LDL receptor has at least three functions: (1) it is necessary for the binding of LDL, but not of β -VLDL, to intact cells; (2) it is necessary for acid-dependent dissociation of β -VLDL; and (3) its presence is necessary to allow the receptor to engage in repeated recycling without degradation. Despite the removal of the EGF precursor region, which is about half the protein, the deleted LDL receptor performed many of its functions normally. It moved from the endoplasmic reticulum to the Golgi complex at a normal rate, as indicated by the rapid processing of O-linked sugars that resulted in a large increase in apparent M_r . The receptor reached the cell surface, as indicated by its ability to bind ^{125}I -labelled IgG-C7 and ^{125}I -labelled β -VLDL. Thus, the protein was not denatured, because denatured membrane proteins do not generally leave the endoplasmic reticulum^{21,22}. The deleted receptor bound ^{125}I -labelled β -VLDL with normal affinity and mediated its internalization at a normal rate.

Although the deleted receptor failed to bind ^{125}I -labelled LDL on the cell surface, it did bind ^{125}I -labelled LDL after SDS gel electrophoresis and transfer to nitrocellulose. This finding suggests that one function of the EGF precursor region is to place the binding domain in a conformation that renders it accessible to LDL on the cell surface. This EGF precursor region does not seem to be required for accessibility to β -VLDL. A similar loss of binding of LDL, but not β -VLDL, was observed previously in a naturally occurring mutant LDL receptor in which one of the seven ligand binding repeats was deleted from the protein²³. The active moiety of LDL, apoB-100, is a large protein of ~514K whose binding sites must have a relatively fixed configuration;

We interpret the data in Fig. 7*a* and *b* to indicate the normal receptors that bind β -VLDL recycle somewhat less efficiently than normal receptors that have bound LDL. In the receptors produced by p Δ EGFP, this effect is markedly accentuated, and the receptors gradually disappear from the surface after incubation with β -VLDL. The receptors that disappear from the surface of the p Δ EGFP-transfected cells are rapidly degraded. This conclusion was reached from studies in which we prelabelled the LDL receptor by incubation with ^{35}S -methionine and then examined the degradation of the labelled receptor on subsequent incubation with or without β -VLDL (Fig. 7, lower panels). Incubation with β -VLDL for up to 6 h had no effect on the degradation of ^{35}S -labelled receptors in cells transfected with the normal cDNA (Fig. 7*c*). In striking contrast, incubation with β -VLDL for 4 h caused the complete degradation of ^{35}S -labelled receptors in the p Δ EGFP-transfected cells (Fig. 7*d*).

Acid-mediated release

In seeking an explanation for the deficient recycling and enhanced degradation of the deleted receptor in the presence of β -VLDL, we tested the ability of acid to release β -VLDL from the two receptors (Fig. 8). The ^{125}I -labelled β -VLDL was allowed to bind to the receptor on the cell surface at 4 °C. The cells were then washed and incubated in buffers at varying pH either at 37 °C (Fig. 8*a*) or at 4 °C (Fig. 8*b*), after which we measured the amount of ^{125}I -labelled β -VLDL that dissociated from the receptor and appeared in the culture medium. At pH 7.5 there was some dissociation from both receptors. When the pH was reduced to 5.5, the normal receptor showed a marked

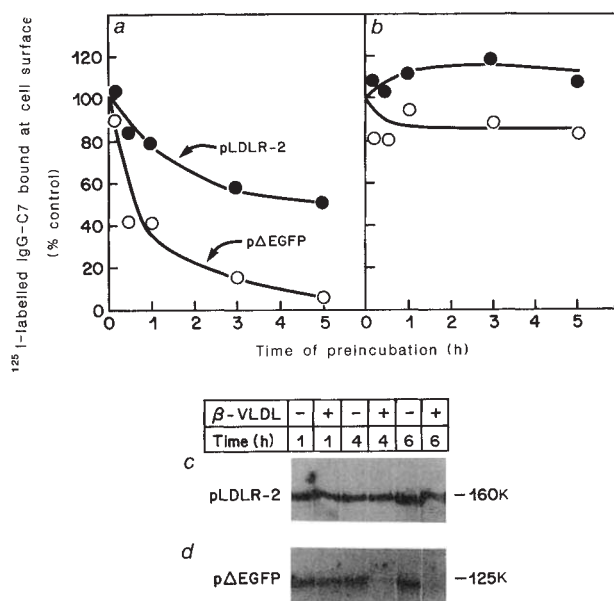


Fig. 7 *a, b*, Decrease in cell-surface LDL receptors in cells transfected with pLDLR2 (●) or pΔEGFP (○) after incubation with β-VLDL. On day 4 of cell growth, duplicate monolayers of cells expressing the indicated cDNA received 2 ml of warm medium D containing either 10 μg protein per ml of unlabelled β-VLDL (*a*) or 20 μg protein per ml of unlabelled LDL (*b*). After preincubation for the indicated time at 37 °C, each dish was rapidly washed twice with ice-cold buffer C and then incubated for 1 h at 4 °C with 10 mM suramin to remove surface bound lipoproteins (see Fig. 5), followed by two rapid washes with ice-cold buffer C. Each monolayer then received 1.5 ml of ice-cold medium B containing 2 μg ml⁻¹ of ¹²⁵I-labelled IgG-C7 (652 c.p.m. ng⁻¹) in the absence or presence of 300 μg ml⁻¹ unlabelled IgG-C7. After incubation for 1 h at 4 °C, the amount of high-affinity binding of ¹²⁵I-labelled IgG-C7 was determined. For the pLDLR2 and pΔEGFP-transfected cells, the "100% of control" values were 213 and 76 ng per mg protein, respectively. *c, d*, Disappearance of ³⁵S-labelled LDL receptors after incubation of pΔEGFP-transfected cells with β-VLDL. On day 2 of growth, cells expressing the indicated cDNA received methionine-free Ham's F-12 medium containing 88 μCi ml⁻¹ ³⁵S-methionine, 10 μM unlabelled methionine, 2 mM glutamine, 5% lipoprotein-deficient serum, 0.1 μg ml⁻¹ 25-hydroxycholesterol, and 10 μg ml⁻¹ cholesterol. After incubation for 16 h at 37 °C, each monolayer was washed twice with phosphate-buffered saline and then incubated at 37 °C for the indicated time in complete medium (0.2 mM methionine) containing either no β-VLDL or 20 μg protein per ml of β-VLDL as indicated. Radio-labelled LDL receptors were precipitated from cell extracts and processed for SDS gel electrophoresis as described in Fig. 2. The dried gels were exposed to XAR-5 film for 18 h at -70 °C.

β-VLDL contains multiple copies of apoE, a 33K protein that is likely to be mobile on the surface of the lipoprotein. ApoE may be able to reach its binding site on mutant LDL receptors through its ability to slide along the lipoprotein surface.

The most striking abnormality of the deleted LDL receptor was its disappearance from the surface after binding β-VLDL. Approximately 50% of the receptors disappeared from the surface within 1 h; 90% disappeared within 5 h (Figs 6 and 7). Within 4 h the receptor was degraded, as indicated by the decline in total immunoprecipitable ³⁵S-methionine-labelled receptors (Fig. 7*d*). Considered together, the data suggest that the deletion-bearing receptor, once having bound β-VLDL, entered the cell rapidly but returned to the surface at a relatively slow rate. Over time an appreciable fraction of the internalized receptors was degraded.

In the presence of β-VLDL the number of receptors declines

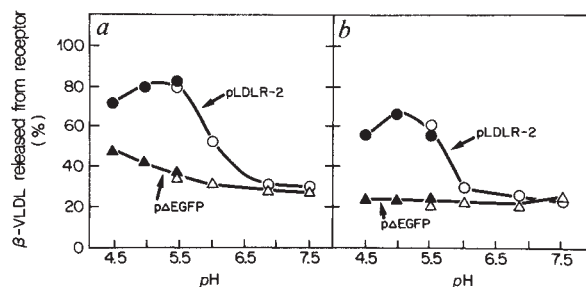


Fig. 8 Acid-dependent release at 37 °C (*a*) or 4 °C (*b*) of receptor-bound ¹²⁵I-labelled β-VLDL from the surface of cells transfected with pLDLR2 (circles) or pΔEGFP (triangles). On day 4 of cell growth, duplicate monolayers of cells transfected with the indicated cDNA received 1.5 ml of ice-cold medium B containing 5 μg protein per ml of ¹²⁵I-labelled β-VLDL (325 c.p.m. per ng protein). After incubation for 2 h at 4 °C, each monolayer was rapidly washed six times with ice-cold buffer C. The dishes then received 3 ml of buffer containing either 50 mM Tris-maleate (○, △) or 50 mM Tris-succinate (●, ▲) at the indicated pH plus 50 mM NaCl, 2 mM CaCl₂, and 5% human lipoprotein-deficient serum. After incubation for either 5 min at 37 °C (*a*) or 15 min at 4 °C (*b*), the amount of ¹²⁵I-labelled β-VLDL released into the buffer was determined by scintillation counting. The cells were dissolved in 0.1 M NaOH and the amount of ¹²⁵I-labelled β-VLDL remaining bound to the cells was measured. The amount of ¹²⁵I-labelled β-VLDL released is expressed as a percentage of the total radioactivity recovered from each dish (buffer plus cell-bound). For the pLDLR2 and pΔEGFP-transfected cells, these values averaged 412 and 341 ng per mg protein, respectively (*a*), and 368 and 379 ng per mg protein, respectively (*b*).

more slowly than would be expected if each receptor simply entered the cell and failed to reappear. In this case 80% of the receptors should have disappeared within 30 min, as estimated from the internalization rate shown in Fig. 5. However, only 50% of the receptors disappeared within the first hour (Figs 6 and 7). It is possible that some of the receptors can return to the surface despite the loss of acid-dependent ligand release. Indeed, the rate of dissociation of β-VLDL from the receptor is appreciable, even without the acid-dependent component (Fig. 8). Alternatively, at the beginning of the experiment the cell might have contained a pool of recycling receptors that had previously entered the cell without ligand. These receptors could continue to reappear on the cell surface after β-VLDL is added, thus replacing the receptors that were sequestered after β-VLDL binding and leading to a slow disappearance of total receptors from the cell surface. Preliminary evidence from electron microscopy suggests that the receptors encoded by pΔEGFP recycle normally in the absence of β-VLDL, as indicated by their localization within coated pits prior to incubation with β-VLDL (unpublished observations).

Surprisingly, the acid-mediated ligand dissociation does not appear to be dependent on a simple titration of charged amino acids in the ligand binding domain. Such a titration should have occurred even in the absence of the EGF precursor region, yet the ligand was not released. Rather the EGF precursor region must undergo some conformational change at acid pH, which in turn must cause the β-VLDL to dissociate from the ligand binding domain. Whether the conformational change is entirely intramolecular, or whether it involves the formation or dissolution of multimeric complexes between the LDL receptor and itself or other proteins, is not known.

Our studies were performed with a receptor in which we removed the entire EGF precursor region, including the long spacer between growth factor repeats B and C (Fig. 1). We have also prepared a construct in which only the A and B repeats were deleted. In preliminary experiments we have found that this receptor also shows a defect in acid-dependent β-VLDL

dissociation and impaired recycling, thus suggesting that these functions have a requirement for the growth factor repeats themselves.

If the present findings can be generalized to other proteins with growth factor repeats, the data would suggest that these repeats help to mediate acid-dependent conformational changes or protein-protein interactions. Insofar as is known, all the protein domains that contain growth factor repeats are extracellular and all these proteins engage in protein-protein interactions, as typified most clearly by the blood clotting factors and the cell surface receptors. Whether all these proteins undergo

acid-dependent conformational changes is not known, but the hypothesis is readily testable.

We thank Edith Womack and Deborah Darrie for assistance with the tissue culture experiments; and Gloria Brunschede, Wen Ling Niu, and Debra Noble-Schlesinger for technical assistance. This research was supported by grants from the NIH and the Lucille P. Markey Charitable Trust. C.G.D. was the recipient of a postdoctoral fellowship from the NIH. T.C.S. was the recipient of a postdoctoral fellowship from the Deutsche Forschungsgemeinschaft. D.W.R. is the recipient of a Research Career Development Award from the NIH.

Received 12 March; accepted 19 March 1987.

1. Doolittle, R. F., Feng, D.-F. & Johnson, M. S. *Nature* **307**, 558-566 (1984).
2. Südhof, T. C. *et al. Science* **228**, 893-895 (1985).
3. Jackman, R. W., Beeler, D. L., VanDeWater, L. & Rosenberg, R. D. *Proc. natn. Acad. Sci. U.S.A.* **83**, 8834-8838 (1986).
4. Dahlback, B., Lundwall, A. & Stenflo, J. *Proc. natn. Acad. Sci. U.S.A.* **83**, 4199-4203 (1986).
5. Wharton, K. A., Johansen, K. M., Xu, T. & Artavanis-Tsakonas, S. *Cell* **43**, 567-581 (1985).
6. Greenwald, I. *Cell* **43**, 583-590 (1985).
7. Doolittle, R. F. *Trends biochem. Sci.* **10**, 233-237 (1985).
8. Südhof, T. C., Goldstein, J. L., Brown, M. S. & Russell, D. W. *Science* **228**, 815-822 (1985).
9. Davis, C. G. *et al. J. biol. Chem.* **261**, 2828-2838 (1986).
10. Goldstein, J. L., Brown, M. S., Anderson, R. G. W., Russell, D. W. & Schneider, W. J. *A. Rev. Cell Biol.* **1**, 1-39 (1985).
11. Brown, M. S., Anderson, R. G. W. & Goldstein, J. L. *Cell* **32**, 663-667 (1983).
12. Helenius, A., Mellman, I., Wall, D. & Hubbard, A. *Trends biochem. Sci.* **8**, 245-250 (1983).
13. Yamamoto, T. *et al. Cell* **39**, 27-38 (1984).
14. Zoller, M. J. & Smith, M. *DNA* **3**, 479-488 (1984).
15. Kingsley, D. M. & Kreiger, M. *Proc. natn. Acad. Sci. U.S.A.* **81**, 5454-5458 (1984).
16. Tolleshaug, H., Hobgood, K. K., Brown, M. S. & Goldstein, J. L. *Cell* **32**, 941-951 (1983).
17. Daniel, T. O., Schneider, W. J., Goldstein, J. L. & Brown, M. S. *J. biol. Chem.* **258**, 4606-4611 (1983).
18. Brown, M. S. & Goldstein, J. L. *J. clin. Invest.* **72**, 743-747 (1983).
19. Mahley, R. W. & Innerarity, T. L. *Biochim. biophys. Acta* **737**, 197-222 (1983).
20. Schneider, W. J., Beisiegel, U., Goldstein, J. L. & Brown, M. S. *J. biol. Chem.* **257**, 2664-2673 (1982).
21. Gething, M.-J., McCammon, K. & Sambrook, J. *Cell* **46**, 939-950 (1986).
22. Yamamoto, T., Bishop, R. W., Brown, M. S., Goldstein, J. L. & Russell, D. W. *Science* **232**, 1230-1237 (1986).
23. Hobbs, H. H., Brown, M. S., Goldstein, J. L. & Russell, D. W. *J. biol. Chem.* **261**, 13114-13120 (1986).
24. Maniatis, T., Fritsch, E. F. & Sambrook, J. *Molecular Cloning, a Laboratory Manual*, 1-545 (Cold Spring Harbor Laboratory Press, New York, 1982).
25. Huettinger, M., Schneider, W. J., Ho, Y. K., Goldstein, J. L. & Brown, M. S. *J. clin. Invest.* **74**, 1017-1026 (1984).
26. Goldstein, J. L., Basu, S. K. & Brown, M. S. *Meth. Enzym.* **98**, 241-260 (1983).
27. Beisiegel, U., Schneider, W. J., Goldstein, J. L., Anderson, R. G. W. & Brown, M. S. *J. biol. Chem.* **256**, 11923-11931 (1981).

Behaviour modification by *in vitro* mutagenesis of a variable region within the period gene of *Drosophila*

Qiang Yu, Hildur V. Colot, Charalambos P. Kyriacou*, Jeffrey C. Hall & Michael Rosbash

Department of Biology, Brandeis University, Waltham, Massachusetts 02254-9110, USA

* Department of Genetics, University of Leicester, University Road, Leicester LE1 7RH, UK

The period gene of Drosophila melanogaster, implicated in the control of both the circadian and male courtship song rhythms, is found to be polymorphic. Alleles differ in the length of a region of the gene encoding a series of threonine-glycine repeat units. The phenotypes of transformed fruit flies, in which the only functional period gene lacks the entire perfect threonine-glycine repeat region, show that the effects of the period gene on the circadian and male courtship song rhythms can be dissociated.

THE *period* (*per*) locus of *Drosophila melanogaster* is one of the best studied 'behavioural genes' in any organism. A considerable amount of cytogenetic and behavioural information on *per* has accumulated over the past 15 years¹⁻³. Recently, the locus has been cloned, the relevant transcript(s) defined and studied, the protein sequence(s) predicted from DNA sequencing studies and analysis of *per* mutants, and arrhythmic *per* mutants rescued with wild-type *per* DNA in P-element-mediated transformation experiments⁴⁻¹¹.

These studies have led to the following picture. The *per* gene is a non-vital locus, as flies in which it is deleted are viable³⁻⁵. However, the absence or inactivity of the gene (in *per*-deleted or *per*⁰ mutants) leads to a phenotype in which the circadian locomotor activity rhythm, the circadian eclosion rhythm, and the rhythm of ~1 min periodicity of the male courtship song are not present^{1,2}. Also, short- and long-period mutations of the gene exist each of which causes all three types of periodicities to be shortened or lengthened, respectively (refs 1 and 2). Because the original arrhythmic, short-period and long-period

mutations have been localized to single nucleotides^{11,12}, it is certain that the gene product is not only essential for normal rhythmicity, but also that modifications of the protein products affect, directly or indirectly, the timing machinery. The *per* gene exerts its influence on these timing phenomena in the nervous system, because both genetic mosaic and molecular experiments indicate that this is the site of *per* gene expression^{13,14}. As expression is evident in the embryo as well as the adult^{14,15}, *per* might affect a particular aspect of nervous system development necessary for proper rhythmic function and/or might have a physiological function in which the gene product interacts more directly with central pacemakers.

Despite this extensive description of the *per* gene and its expression, its biochemical function is unknown. The DNA sequence and predicted protein sequence contain few clues, except for a stretch of alternating threonines and glycines^{9,16}. This Thr-Gly repeat region is homologous to mouse DNA potentially coding for Thr-Gly and Ser-Gly repeats¹⁷. That it is also similar to a Ser-Gly repeat region in a rat chondroitin

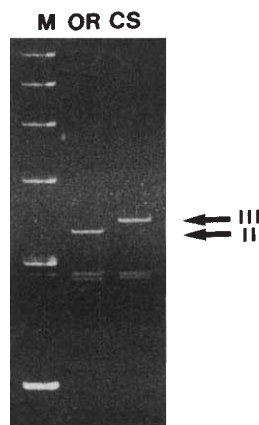


Fig. 1 The Thr-Gly fragment from Oregon-R and Canton-S DNA. The *NcoI*-*Bam*HI fragment (positions 4,403 to 5,627)¹⁰ was purified from a low melting temperature agarose gel, digested with *Cfo*I (Boehringer Mannheim), run on a 7.5% polyacrylamide gel, and stained with ethidium bromide. The arrows indicate the TG-repeat-containing fragments corresponding to Type II and III (303 and 321 bp, respectively; see text). M, 123-bp size marker ladder (BRL); OR, Oregon-R; CS, Canton-S (cloned from the *per*^s mutant¹¹, which was derived from Canton-S). The flanking fragments shown are 242, 237 and 147 bp.

sulphate proteoglycan¹⁸ is the basis for the suggestion that *per* may code for a proteoglycan^{9,16}. In one case, preliminary biochemical data supported this inference¹⁶.

As the threonines in the Thr-Gly repeat are the predicted sites of attachment of carbohydrate side chains¹⁸, we sought to test the importance of this region in two ways. First, we examined the evolutionary conservation of this repeated sequence by comparing different strains of *D. melanogaster*. Second, we examined the functional significance of this region by constructing a *per* gene deletion derivative lacking all the perfect Thr-Gly repeat. The first approach revealed a surprising degree of polymorphism and evolutionary plasticity in the predicted protein sequence of the Thr-Gly repeat region. The second approach indicated that the Thr-Gly region may have a more important function in the courtship song phenotype than in the locomotor activity (circadian) phenotype. When considered together, our data suggest that the repeated DNA sequence in the *per* coding region 'creates' a considerable degree of evolutionary divergence. This divergence at the DNA and protein level may be significant for the evolutionarily variable courtship song rhythm, but may influence the more conserved circadian rhythm only marginally.

Evolutionary comparisons

In comparing our sequence of the *per* gene from the Oregon-R wild-type strain of *Drosophila melanogaster*⁵ with that of the *per* gene from the Canton-S strain⁹, we noticed that the Thr-Gly repeat region was three repeat units, or 18 nucleotides, longer in Canton-S than in Oregon-R. To ensure that this difference was genuine, we compared restriction digests of a portion of Canton-S and Oregon-R *per* DNA by acrylamide gel electrophoresis (Fig. 1). Subclones containing the Thr-Gly repeat from both strains were digested with *Cfo*I, a restriction enzyme that generates small fragments and for which, as shown by sequence analysis, there are no sites in the Thr-Gly repeat region. DNA from each strain gives rise to a Thr-Gly-containing fragment; the one from Oregon-R clones is indeed about 18 nucleotides (nt) shorter than that from Canton-S (Fig. 1), as predicted from the sequences.

The only remaining possible artefact was that the size difference was generated during cloning in *Escherichia coli*. To test this possibility and to examine additional strains for the

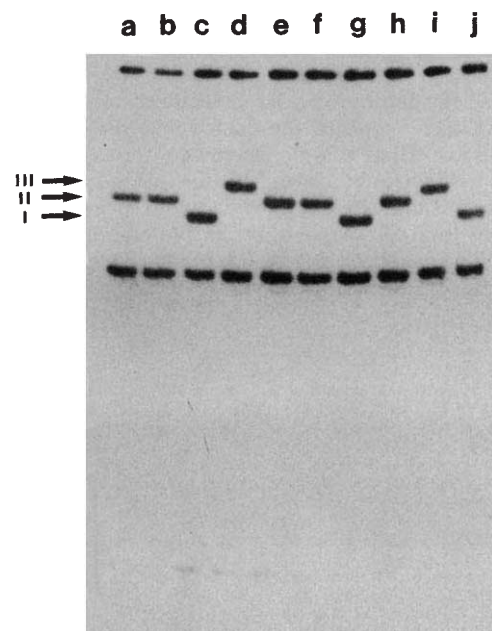


Fig. 2 Variability of the TG-repeat region among different strains of *Drosophila melanogaster*. The TG-containing bands corresponding to classes I, II and III (see text) are indicated by arrows; they are 285, 303 and 321 bp long, respectively. The flanking fragments shown are 570, 237+232 and 94 bp long. Lanes a-j: DNA from various strains of *D. melanogaster*, mostly obtained from W.M. Gelbart; a, Seto; b, 'π 2'; c, Oregon-R (from R. J. Konopka); d, Canton-S (from M. W. Young); e, '+6'; f, Harwich; g, '78.25'; h, *ci*^D/*ey*^D; i, OK1; j, Chieti V. (Note that the 'Oregon-R' strain used here is apparently different from the one shown in Fig. 1 and discussed in the text.)

Methods. DNA was prepared from 10 flies of each strain as described²⁴. The entire pellet (~3 µg) was digested with 60 units of *Cfo*I for 4 h, followed by RNase digestion (80 µg ml⁻¹) for 25 min at 37 °C, phenol/chloroform extraction, and precipitation with ethanol in the presence of 2.5 M ammonium acetate. The entire digest was run on a denaturing 4% acrylamide, 8 M urea gel in 1×TBE, and the gel was electroblotted onto Biotrans (ICN) in 0.5×TBE at 450 mA for 45 min. The blot was UV-treated (800 µW cm⁻²) for 2 min, baked for 1 h, prehybridized overnight and hybridized with probe at 55 °C under standard conditions. The probe was an SP6 transcript of the *Nco*I-*Bam*HI fragment (positions 4,403 to 5,627)¹⁰.

size of their Thr-Gly regions, whole-genome Southern blot analysis was performed (Fig. 2). DNA preparations from 11 strains were digested as in Fig. 1, the fragments separated by acrylamide gel electrophoresis and the DNA electroblotted to a nylon membrane. Hybridization with a Thr-Gly-repeat-containing probe revealed a third size class, apparently ~36 nt shorter than that from Canton-S DNA. In fact, inspection of these 11 strains (Fig. 2) and 19 others (data not shown) suggests that each of the 30 strains fall into one of three size classes: two strains in the largest class (III), 16 strains in the middle class (II), and 12 strains in the smallest class (I) (see Fig. 2).

As examples of the large and medium classes had already been sequenced, we cloned an example of the smallest fragment and subjected it to DNA sequence analysis. It was in fact 18 nt shorter than the second class (see below). Thus, the Thr-Gly repeat region of *D. melanogaster* consists of (at least) three size classes: 17 pairs of perfect tandem repeats, 20 pairs and 23 pairs (Fig. 3).

Although the repeat region contains no discontinuities at the amino-acid level (merely alternating threonines and glycines), the third positions of the codons provide a clue as to where,

region has a much greater impact on the courtship-song cycle durations than on the circadian rhythm frequency.

Implications of *per* variability

The inter-strain variability in the Thr-Gly repeat region is particularly noteworthy because it is unusual. In a comparison of more than 5,000 base pairs (bp) of the *per* gene from Canton-S⁹ and Oregon-R¹⁰, no other insertions or deletions are evident. Only 14 substitutions are visible; five are in the introns, and nine are third base, isocoding substitutions. Preliminary data also exist for the *D. simulans per* gene. A comparison of 2.4-kb of sequenced DNA with the corresponding region of the *D. melanogaster per* gene reinforces and strengthens this conclusion, as multiple insertion-deletion events are evident in the Thr-Gly repeat region and its environs (H.V.C., unpublished data and A. C. Jacquier and D. A. Wheeler, personal communication).

Although the variation we observe in the *per* Thr-Gly coding region might simply reflect a lack of selective constraints on the region relative to the rest of the gene, in our view it is more likely that the sequence of the region itself makes insertion and deletion events more frequent. A possible mechanism for this has in fact been suggested¹⁹, and involved the 'catalysis' of unequal genetic exchanges and/or replication errors by repeated sequences. Insertions and deletions of multiples of the Thr-Gly repeat unit maintain an open reading frame in the *per* gene; this would again make it more likely for the variants so generated to become fixed in the population as compared to deletions or insertions elsewhere in the gene, two out of three of which would cause a frameshift. Our observation that the 18-bp perfect repeat is the unit of insertion or deletion fits well with this scheme, presumably indicating that the sequence is a good substrate for the events that generate the *per* polymorphism.

Implications of *in vitro* mutagenesis of *per*

To test the function of this region we generated a deletion of the entire perfect Thr-Gly repeat. Although somewhat extreme,

we followed this experimental strategy because the behavioural phenotypes of the transformants generated to date have, in many cases, been non-wild-type and rather variable^{6,8,10,11}. Thus, we thought it might be difficult to detect meaningful behavioural differences, such as might be induced by small changes in the size of the Thr-Gly repeat unit.

The behavioural effects of the substantial change we engineered are surprising. Deletion of the Thr-Gly repeat has little or no effect on the circadian phenotype of the transformed flies, although more subtle effects cannot be excluded by the experiments reported here. These findings are analogous to those from experiments on the LDL (low-density lipoprotein) receptor gene, in which deletion of a region coding for a high percentage of serines and threonines—known to be modified by O-linked glycosylation—had no detectable effect on LDL receptor function²⁰. Although we do not have enough biochemical information to interpret our observation clearly, two obvious possibilities present themselves. First, the circadian functions of the *per* protein(s) may take effect through a non-proteoglycan route, insensitive to the absence of the Thr-Gly region^{9,16,18}. Second, other Thr-Gly or Ser-Gly pairs elsewhere in the *per* protein sequence may carry GAG (glycosamino-glycan) side chains, and these modifications may be sufficient for circadian functions of *per*. (Note that there are three Thr-Gly pairs directly downstream of the perfect repeat which are not removed by the deletion.)

Whereas the circadian results alone lead to the suggestion that the Thr-Gly repeat region is of little or no importance to fly behaviour, the profound effect of the deletion on the courtship song belies this interpretation. The Thr-Gly repeat region and the carbohydrate side chains it may carry may specifically or preferentially influence the quantitative characteristics of the courtship song rhythm. Alternatively, the deletion may change the structure of the *per* protein(s) in another way that affects only its courtship song function.

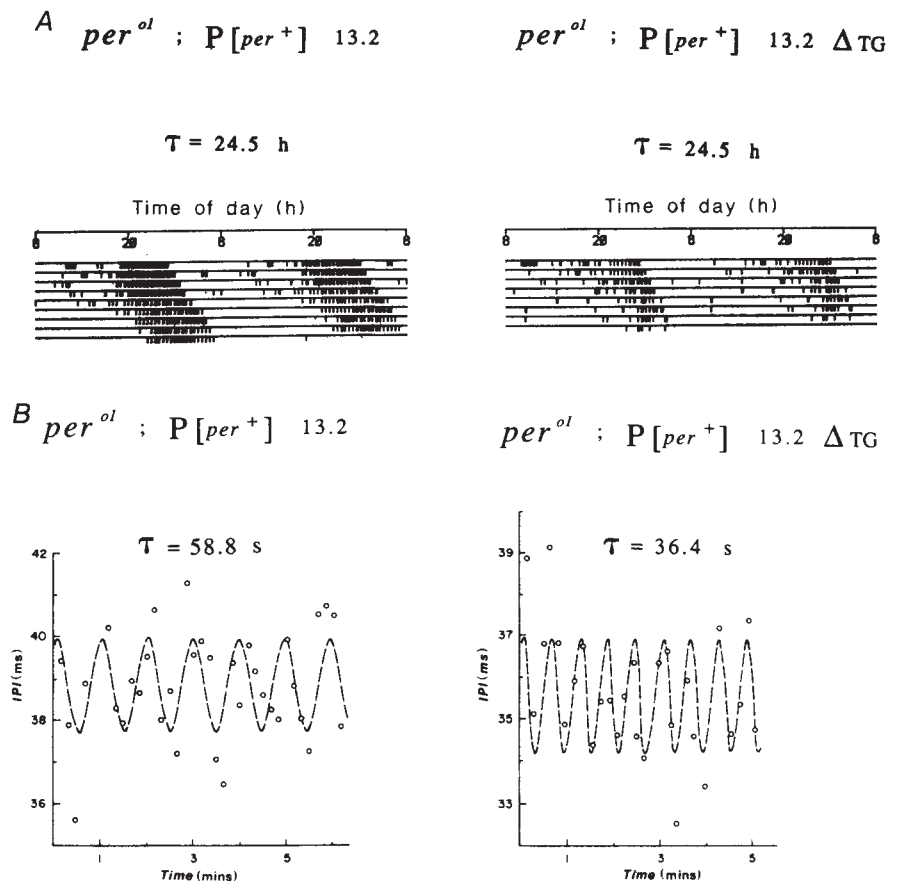
In any case, this deletion mutant, in contrast to the *per* mutants described previously, separates the effect of the *per* gene on the

Table 1 Circadian rhythms and courtship song rhythms of transformants

Genotype	Circadian rhythms			Courtship song rhythms		
	Rhythmic (n)	Mean period ±s.e.m. (h)	Arrhythmic (n)	Rhythmic (n)	Mean period ±s.e.m. (s)	Arrhythmic (n)
Controls						
<i>per</i> ⁺	13	24.1 ± 0.1	0	5*	59.0 ± 3.7	2
<i>per</i> ⁰¹	0	—	20	0*	—	7
Transformants						
13.2	32	24.6 ± 0.1	1	5	59.5 ± 4.5	1
13.2 ΔTG	28	24.9 ± 0.1	1	7	40.4 ± 4.1	1
8.0	17	25.5 ± 0.4	6	19	71.0 ± 3.8	18
8.0 ΔTG	12	25.9 ± 0.7	9	13	46.4 ± 4.2	12

Germline transformants, whose genetic background was *per*⁰¹; *ry*⁵⁰⁶, were produced as described previously^{8,11}, with transformed lines being initiated from individual flies expressing the *ry*⁺ marker carried in the P-sequence-containing vector (see Fig. 4). By behaviourally testing *ry*⁺ individuals from the established lines, circadian locomotor activity and courtship song data were collected and analysed as described previously^{2,8} (see also Fig. 5 legend). The control results for song rhythms (indicated by *), from recently analysed normal and arrhythmic *per* mutant males, are from ref. 8. Both transformant and control males were analysed 'blind', with decisions on rhythm periods being made before a given song record was 'decoded' as to the genotype involved. Our controls, involving determinations of circadian periods for separately tested groups of *per*⁺ *D. melanogaster*, have given values ranging from 23.9 to 24.1 h (additional data in refs 6 and 8). The transformant types are named according to the kind of *per*-derived DNA insert they carried: a 13.2-kb genomic fragment¹¹ (two independently isolated strains); a 13.2-kb fragment from which the Thr-Gly repeat had been removed (ΔTG, see Fig. 4; five strains); an 8.0-kb fragment¹¹ (two strains); and an 8.0-kb fragment with the TG repeat deleted (three strains). Each transformed individual tested was a male carrying one dose of a transduced *per*-derived insert (see Fig. 5 legend) on an autosome. Data from within a transformant type (designated for example, '8.0' or '8.0ΔTG') were statistically indistinguishable and thus were pooled. For the comparison of the 13.2 and 13.2 ΔTG types, the small difference between the mean circadian periods is significant by one-way ANOVA ($P < 0.02$), as are the song rhythm periods ($P < 0.05$). The slightly longer circadian periodicity associated with the 13.2ΔTG strains may be an artefact of our having tested most of the 28 flies from a particular strain whose mean period was 25.2 ± 0.2 h ($N = 11$). Although this value was not significantly different from that for the other four strains of this type, given the relatively small numbers involved, further data may eventually reveal that it is meaningfully longer, perhaps because of a 'position effect' associated with the location of this particular insert. Note, however, that we have not detected 'major' position effects for our many types of transformants, that is, between transformants receiving a given category of *per*-locus derived DNA fragment^{6,8,11}. For comparison of the 8.0 and 8.0ΔTG types, the circadian periods are indistinguishable ($P > 0.6$), whereas the song periodicities are much shorter in the ΔTG males ($P \leq 0.01$).

Fig. 5 Circadian and song rhythms of transformants. **A**, Locomotor activity data of adult males were collected and plotted as in ref. 8. The records here are 'double plots' of free-running activity collected in constant darkness, after four days of entrainment by 12 h:12 h light:dark cycles (days 1-2 of free-running activity plotted horizontally on the first line, days 2-3 on the second line, and so on). The best estimates of the circadian periods (τ) were determined from Chi-square periodogram analyses⁸. Each transformed individual was hemizygous for *per*^{ol} (as shown) and also homozygous for the *ry*⁵⁰⁶ marker on chromosome 3. The *ry*⁺ insert for the fly in the left-hand panel is the 13.2-kb *per*-locus genomic fragment, carried in one dose (that is, a *per*⁺ insert heterozygous with a non-insert-bearing chromosome). The transformant whose behaviour is in the right-hand panel carried one dose of a *ry*⁺, TG-deleted insert. **B**, Courtship song 'pulses' were recorded (as in refs 2 and 10) from individual males, each of which courted a female off and on for 5-6 min. The genotypes of the two males here are as described in **A**. Mean interpulse intervals (IPIs) were computed for a series of 10-s time frames, then plotted and analysed as in ref. 2, whereby the best-fitting sine wave, defining the rhythmic IPI fluctuations, was determined eventually by an *F*-ratio test (a probability of 5% or less was demanded for a significant fit). The periods (τ) for these two examples were taken from the parameters associated with the best fits. Note that the mean IPI points falling 'off' the curves in several instances are nevertheless, in general, located above or below a peak or a trough of the sine wave; this accentuates the chances of data like these being fitted significantly with respect to a sine function and is a further indication of the regularity of these IPI fluctuations.



circadian and courtship song phenotypes. We interpret this observation as indicating that some aspects of the mysterious functions of the *per* gene may be different in their contributions to the circadian and courtship song phenotype. Alternatively, the circadian clock may be well 'buffered' by some feedback mechanism not present, or less active, in the song clock, so that the effects of the mutant (partially deleted) *per* protein(s) are less evident at the level of clock output.

The observations in this report lead to a rather speculative picture which relates repeated DNA to behaviour on the one hand and evolution on the other. The studies of Jeffreys and co-workers¹⁹, and others²¹, provide a framework within which one can understand how 'simple' and repeated DNA sequences, present within the coding region of a gene, could be a source of intra-strain polymorphism and inter-strain difference at the amino-acid level. The results presented here reinforce this possibility. It is noteworthy that this variability is present in a gene with behavioural consequences, as it is precisely in genes of this

kind where such variability might be put to good use. Indeed, it is the behaviourally relevant^{22,23} and variable courtship song phenotype that is most affected by this region and the circadian phenotype, which is essentially identical interspecifically (see Table 1 caption), that is least affected. Thus it may be that the variable length of the Thr-Gly repeat is a fundamental source of behavioural variability. As the courtship song may be involved in mate selection²², it is even possible that the variability in the Thr-Gly repeat region DNA, which leads to variation in the *per* protein sequences, contributes to essential behaviour mechanisms involved in critical aspects of reproductive isolation.

We thank our colleagues at Brandeis University for helpful comments on these experiments and on the manuscript. We also thank G. Miklos for his comments on the manuscript and D. Wheeler and A. C. Jacquier for communication of unpublished results. This work was supported by an NIH grant to M.R. and J.C.H.

Received 5 January; accepted 9 March 1987.

- Konopka, R. J. & Benzer, S. *Proc. natn. Acad. Sci. U.S.A.* **68**, 2112-2116 (1971).
- Kyriacou, C. P. & Hall, J. C. *Proc. natn. Acad. Sci. U.S.A.* **77**, 6929-6933 (1980).
- Smith, R. F. & Konopka, R. J. *Molec. gen. Genet.* **185**, 30-36 (1982).
- Bargiello, T. A. & Young, M. W. *Proc. natn. Acad. Sci. U.S.A.* **81**, 2142-2146 (1984).
- Reddy, P. *et al. Cell* **38**, 701-710 (1984).
- Zehring, W. A. *et al. Cell* **39**, 369-376 (1984).
- Bargiello, T. A., Jackson, F. R. & Young, M. W. *Nature* **312**, 752-754 (1984).
- Hamblen, M. *et al. J. Neurogenet.* **3**, 249-291 (1986).
- Jackson, F. R., Bargiello, T. A., Yun, S.-H. & Young, M. N. *Nature* **320**, 185-188 (1986).
- Citri, Y. *et al. Nature* **326**, 42-47 (1987).
- Yu, Q. *et al. Proc. natn. Acad. Sci. U.S.A.* **84**, 784-788 (1987).
- Baylies, M. K., Bargiello, T. A., Jackson, F. R. & Young, M. W. *Nature* **326**, 390-392 (1987).

- Konopka, R. J., Wells, S. & Lee, T. *Molec. gen. Genet.* **190**, 284-288 (1983).
- James, A. A., Ewer, J., Reddy, P., Hall, J. C. & Rosbash, M. *EMBO J.* **5**, 2313-2320 (1986).
- Young, M. W., Jackson, F. R., Stein, H.-S. & Bargiello, T. A. *Cold Spring Harb. Symp. quant. Biol.* **50**, 865-875 (1985).
- Reddy, P., Jacquier, A. C., Abovich, N., Petersen, G. & Rosbash, M. *Cell* **46**, 53-61 (1986).
- Shin, H. S., Bargiello, T. A., Clark, B. T., Jackson, F. R. & Young, M. W. *Nature* **317**, 445-448 (1985).
- Bourdon, M. A., Oldberg, A., Pierschbacher, M. & Ruoslahti, E. *Proc. natn. Acad. Sci. U.S.A.* **82**, 1321-1325 (1985).
- Jeffreys, A. J., Wilson, V. & Thein, S. L. *Nature* **314**, 67-73 (1985).
- Davis, C. G. *et al. J. biol. Chem.* **261**, 2828-2838 (1986).
- Tautz, D., Trick, M. & Dover, G. A. *Nature* **322**, 652-656 (1986).
- Kyriacou, C. P. & Hall, J. C. *Anim. Behav.* **30**, 794-801 (1982).
- Kyriacou, C. P. & Hall, J. C. *Nature* **308**, 62-65 (1984).
- Bender, W., Spierer, P. & Hogness, D. S. *J. molec. Biol.* **168**, 17-33 (1983).

The interstellar spectrum towards the supernova 1987A in the Large Magellanic Cloud

Paola Andreani*‡, Roger Ferlet†
& Alfred Vidal-Madjar†

* Department of Physics, University of Rome, 00185 Rome, Italy

† Institut d'Astrophysique de Paris, 98 bis, Boulevard Arago, 75014-Paris, France

‡ Present address: Institut d'Astrophysique de Paris, 98 bis, Boulevard Arago, 75014-Paris, France

In recent years, dramatic improvements in spectroscopic capabilities have opened up the interstellar line research to the brightest extragalactic targets, and in particular to stars in the Magellanic Clouds (in the visible, see for example refs 1–3). This allows us to tackle two questions of crucial importance: (1) the existence, structure and composition of a gaseous halo surrounding our Galaxy and/or the Clouds; and (2) the fragmentation of the Clouds, their depth along the line of sight, and to probe the interstellar matter in galaxies less evolved than ours. The supernova 1987A offers a unique opportunity to probe the visible interstellar absorption spectrum towards the Large Magellanic Cloud at high resolution and high signal-to-noise ratio. Here we report the first results of these observations. A Ca II 'forest' is detected through all the velocity range between the Galaxy and the Large Magellanic Cloud and in particular, a component is seen at a velocity of $\sim 215 \text{ km s}^{-1}$ which may correspond to cooler gas in a halo. A preliminary abundance study is also presented.

The supernova 1987A, which was discovered on 24 February in the Large Magellanic Cloud (LMC), offers the extraordinary unique opportunity to extend these studies with very high spectral resolution and signal-to-noise ratio observations.

Observations were carried out at the European Southern Observatory in La Silla, Chile during the first part of the nights 24–26 February 1987, with the Coudé Echelle Spectrometer (CES) fed by a 1.4-m telescope and equipped with a cooled Reticon detector of 1,870 pixels, 15- μm each. The wavelength ranges covered are those of (by sample of 30 to 60 Å): in the blue, both lines of the calcium II doublet at 3,933 Å (K) and 3,968 Å (H) and of the ions Ca I and CO^+ around 4,238 Å; and in the red, the sodium-D doublet at 5,893 Å and the potassium I doublet around 7,682 Å. Because the visual magnitude of the supernova was 4.6, the instrumental resolution (full width at half maximum, FWHM), as measured on narrow thorium emission lines and absorption telluric water vapour lines, was set to 3 km s^{-1} ; the internal accuracy of the wavelength calibra-

tion is always better than $\pm 0.2 \text{ km s}^{-1}$. Details of the instrumentation, observational and data reduction procedures can be found in ref. 4.

The final plot for Na I is shown in Fig. 1. It is well known that this wavelength region is crowded with telluric water lines. Following the method developed in ref. 5, we present in Fig. 2 the same enlarged spectrum as in Fig. 1, but after division by the bright template star α Eri; all telluric lines are thus removed. For comparison, the two D-lines have been superimposed on a common heliocentric velocity scale, after correction for the Earth motion. The integration time was one hour; the resulting signal-to-noise ratio (r.m.s.) is about 250, which leads to a minimum detectable equivalent width $W \leq 1 \text{ mÅ}$ (3σ), if we assume that the absorption line widths are of the order of (or smaller than) the instrumental function.

The final enlarged plots for the Ca II H and K lines are shown superposed in Fig. 3. The integration time was only twenty minutes, resulting in a detectable equivalent width of about 1.2 mÅ (3σ).

The region around K I is crowded with atmospheric oxygen absorption lines. After division by a template star to remove these O_2 lines, the limiting detectable equivalent width is about 1.3 mÅ (3σ). Finally, the spectrum in the region of Ca I and CO^+ shows no obvious feature, at a level of about 1 mÅ (3σ).

Figure 3 reveals a completely unknown 'forest' of Ca II absorptions, all over the velocity range between the interstellar material from the disk of the Milky Way expected to show up at around 0 km s^{-1} , and the LMC interstellar gas which should be detected around the systemic velocity of the LMC ($\sim +280 \text{ km s}^{-1}$). No less than 24 components can be distinguished, which are seen in both H and K lines, in their expected ratio of 2 for the unsaturated ones. The locations of these 24 components have been marked on the Na I spectra of Fig. 2. This shows the extremely good correspondence between the Ca II and Na I velocities, but also that fewer components are detected in neutral sodium in spite of a better signal-to-noise ratio. We summarize in Table 1 the velocity centroids and equivalent widths W (accurate to better than $\pm 10\%$ because of the very well-defined flat continua) of all Ca II, Na I and K I interstellar absorption components.

According to the velocities and to the sodium to calcium ratio defined as $x = W(\text{Na I D}_2)/W(\text{Ca II K})$, we can divide the observed components (Table 1) into three main groups that we will briefly discuss separately.

Consider first, galactic gas with $V_\odot \leq 25 \text{ km s}^{-1}$ and $x \geq 2$. The low velocity galactic disk absorption is clearly resolved into three components, with only the middle one being slightly redshifted by about 3 km s^{-1} in Na I. With the exception of this middle one, the other two were already seen at very similar velocities towards two LMC supergiants (R 127 and R 128), observed in Na I, with a resolution of 5 km s^{-1} (ref. 2). The

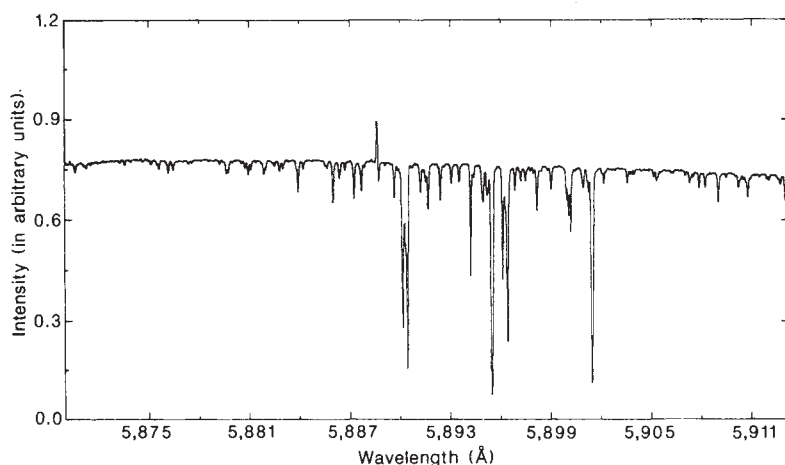


Fig. 1 A one hour Reticon spectrum of the supernova 1987A recorded in the region of the neutral sodium at the ESO-La Silla Observatory on 26 February 1987.

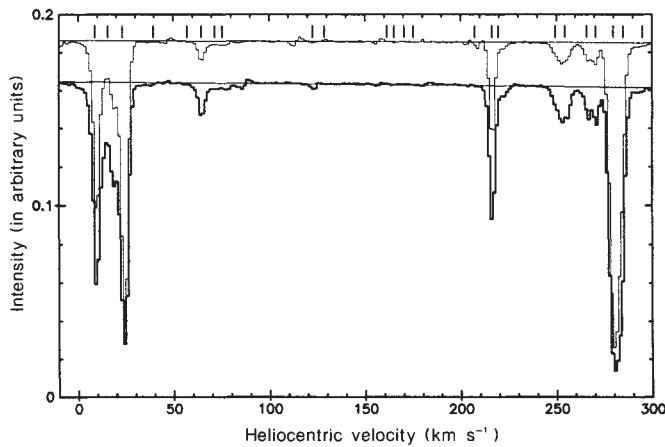


Fig. 2 Same as Fig. 1, after division by a template star in order to release all the telluric water vapour absorption lines present in Fig. 1. The Na I-D₂ (thick line) and D₁ (thin line) lines have been superposed on a common heliocentric velocity scale. At the top, tick marks represent the positions of the 24 absorption components detected in Ca II (Fig. 3). Less components are seen in Na I.

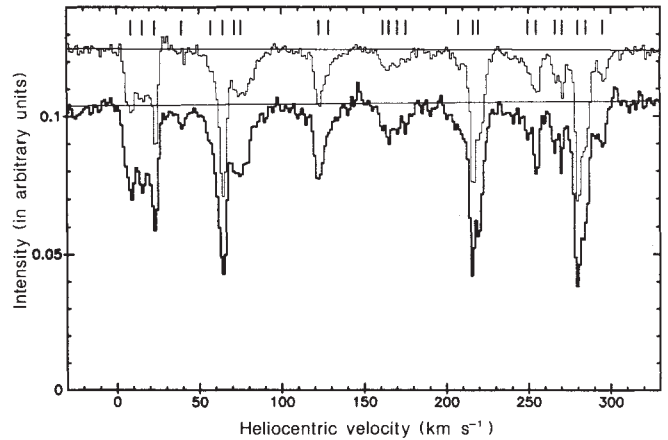


Fig. 3 The Ca II interstellar spectrum towards the supernova 1987A recorded at ESO on 25 February 1987. The K (thick line) and H (thin line) lines are superposed on a common heliocentric velocity scale, and the 24 detected components are marked at the top and described in Table 1.

angular separation between R 127 and R 128 and the supernova is about 1°. Therefore, if we assume a rough distance of 100 pc to this galactic gas, we can infer ~2 pc as an estimation of the size of the interstellar diffuse clouds over which they remain homogeneous in velocity.

The x -value ≈ 2.2 is typical of the galactic disk. The Na I-D lines present saturation effects (doublet ratio < 2) which prevents us from estimates reliable column-densities, unless profile fitting is used.

The only component detected in K I is unsaturated. By using the relation between W and the column-density N corresponding to the linear part of the curve of growth, we derive $N(\text{K I}) = 3.9 \times 10^{10} \text{ cm}^{-2}$.

Second, consider LMC gas with $235 \leq V_{\odot} \leq 280 \text{ km s}^{-1}$ and $1 \leq x \leq 2$. Detailed kinematic comparisons are hardly feasible with existing International Ultraviolet Explorer observations^{6,7}

because of their much lower resolution and accuracy. In the visible region some comparisons can be made, including R 127, R 128, R 139, R 145 and R 136, the central object of the 30 Doradus nebula. It turns out that the 7 Ca II components nos 18 to 24 (Table 1) represent the resolved components of the features seen towards stars located within or around the 30 Doradus nebula, except perhaps for a slightly lower extension towards high velocities. Therefore, one can claim that supernova 1987a is located in or behind the LMC, at least at the distance of the 30 Doradus nebula.

As in the case of the galactic gas, the total equivalent widths measured for the LMC components are somewhat smaller in general than those so far published towards LMC stars. This reflects partly the much lower accuracy of the previous results as well as the higher resolution in conjunction with the clean instrumental profile used here. It could also partly indicate that the supernova is not located entirely behind the 30 Doradus nebula.

Table 1 Interstellar absorption by Ca II, Na I and K I towards the supernova 1987a

Components no.	Ca II-k		Na I-D ₂		Na I-D ₁		K I (7698)	
	$V_{\odot}$ (km s ⁻¹)	W (m Å)	$V_{\odot}$ (km s ⁻¹)	W (m Å)	$V_{\odot}$ (km s ⁻¹)	W (m Å)	$V_{\odot}$ (km s ⁻¹)	W (m Å)
1	8.7	83	9.5	185	9.1	132	21.9	7
2	15.6		18.4		18.7			
3	23.1		23.9		24.2			
4	39.5	4						
5	57.1	70		14	64.5	7		
6	64.6		64.6		71.7			
7	71.6		71.9		77.6			
8	75.4	20	77.8	2		1.0		
9	122.6	28						
10	128.8							
11	161.5							
12	165.4	18						
13	170.4							
14	175.4							
15	207.5	70		38	216.3	21		
16	216.3		216.2		220.6			
17	219.6		222.1					
18	249.6	20		21	252.8	11		
19	254.7		252.7		266.1			
20	266.0		266.4		270.6			
21	270.4	16	270.6	19	280.6	10		
22	279.8		280.6					
23	284.8							
24	294.9	82	293.0	150	289.8	123	279.4	22

There is no saturation effect for components 18, 19 and 20, 21. By using again the linear part of the curve of growth, we derive:

$$\begin{aligned} N(\text{Ca II})_{18,19} &\sim 2.1 \times 10^{11} \text{ cm}^{-2} \\ N(\text{Ca II})_{20,21} &\sim 1.7 \times 10^{11} \text{ cm}^{-2} \\ N(\text{Na I})_{18,19} &\sim 1.1 \times 10^{11} \text{ cm}^{-2} \\ N(\text{Na I})_{20,21} &\sim 1.0 \times 10^{11} \text{ cm}^{-2} \end{aligned}$$

Assuming the H I (21 cm) column-densities given in ref. 8 for two components at $+247 \text{ km s}^{-1}$ and $+268 \text{ km s}^{-1}$ which match very well with the present 18, 19 and 20, 21 Ca II ones respectively, we deduce the ratio $N(\text{Ca II})/N(\text{H I})$ around 2.5×10^{-10} and $N(\text{Na I})/N(\text{H I})$ around 1.2×10^{-10} for the 18, 19 components and, respectively, 1.8×10^{-10} and 1.0×10^{-10} for the 20, 21 components.

The calcium to hydrogen ratio is to be compared to about 18×10^{-10} for the interstellar gas in the solar neighbourhood⁹, but this value is highly dependent on calcium depletion on to grains. The sodium to hydrogen ratio for the corresponding galactic gas is $\sim 13 \times 10^{-10}$ (ref. 10). Finally, the x -values for these components are of the order of unity, somewhat lower than the galactic one of 2.2. All these facts may be an indication of the overall underabundance of metals in the LMC with respect to the Milky Way, along with a smaller depletion of calcium onto grains in the probed regions of the LMC. The only component detected in K I has $N(\text{K I}) \approx 1.2 \times 10^{11} \text{ cm}^{-2}$.

Third, consider all other components with $25 < V_{\odot} < 235 \text{ km s}^{-1}$ and $x < 1$. Although the existence of intermediate-velocity gas is already known from ultraviolet data (refs 6, 7), we have here much more information on kinematics and column-densities. This population is clearly different from the rest because its x -value is below unity.

In the high resolution H I 21 cm data towards 30 Doradus (ref. 8), the Ca II components nos 7, 8 and 9, 10 are clearly detected at the same velocity. For the 7, 8 components, we derive a safe $N(\text{Na I}) = 1 \times 10^{10} \text{ cm}^{-2}$ and a rough estimate of $N(\text{Ca II}) \sim 2 \times 10^{11} \text{ cm}^{-2}$. For 9, 10, only Ca II is detected: $N(\text{Ca II}) \sim 2.9 \times 10^{11} \text{ cm}^{-2}$. Then, we deduce $N(\text{Na I})/N(\text{H I}) \approx$

1.4×10^{-8} and $N(\text{Ca II})/N(\text{H I}) \approx 2.8 \times 10^{-7}$ for the group 7, 8 and $N(\text{Ca II})/N(\text{H I}) \approx 1.7 \times 10^{-7}$ for the group 9, 10. Now the sodium to hydrogen ratio is in good agreement with the galactic linear relation¹⁰, while the calcium to hydrogen ratio is about one hundred times larger than the galactic one. Together with $x < 1$, all this is strongly suggestive of an almost complete absence of depletion.

Because of the galactic latitude of the LMC, the line of sight to the supernova extends mainly through regions of haloes. However, it is still not clear if the intermediate-velocity gas is a signature of the galactic and/or LMC haloes. The velocity analysis towards a rather large sample of Magellanic stars led Songaila *et al.*¹¹ to locate this gas within the LMC. The widely differing velocity and column-density seen simultaneously towards the supernova as discrete components suggests in effect that the LMC is partly made of individual sheets layered over one another. One can note in Fig. 3, that these components seem to be clustered in velocity. A very accurate LMC survey of stellar radial velocities¹² seems to show different groups at velocities above $\sim 200 \text{ km s}^{-1}$ including a few stars around 210 km s^{-1} . The present abundance study is consistent with that interpretation.

This research was supported by ESO and its General Director L. Woltjer, and we wish to thank the ESO staff at La Silla, in particular S. Cristiani, D. Hoffstadt and H. E. Schuster for their help. Our thanks are also due to K. Sahu for having taken low resolution spectra, to M. Péron who helped us with the data processing and to M. C. Pelletan for editing this report.

Received 9 March; accepted 30 March 1987.

1. Songaila, A. *Astrophys. J.* **248**, 945-955 (1981).
2. Ferlet, R., Dennefeld, M. & Maurice, E. *Astr. Astrophys.* **152**, 151-156 (1985).
3. Blades, J. C. *Mon. Not. R. astr. Soc.* **190**, 33-42 (1980).
4. Ferlet, R. & Dennefeld, M. *Astr. Astrophys.* **138**, 303-310 (1984).
5. Vidal-Madjar, A., Ferlet, R., Gry, C. & Lallement, R. *Astr. Astrophys.* **155**, 407-412 (1986).
6. Savage, B. D. & de Boer, K. S. *Astrophys. J.* **230**, L77 (1979).
7. Savage, B. D. & de Boer, K. S. *Astrophys. J.* **243**, 460-484 (1981).
8. McGee, R. X., Newton, L. M. & Morton, D. C. *Mon. Not. R. astr. Soc.* **205**, 1191-1205 (1983).
9. Hobbs, L. M. *Astrophys. J.* **191**, 381-393 (1974).
10. Ferlet, R., Vidal-Madjar, A. & Gry, C. *Astrophys. J.* **298**, 838-843 (1985).
11. Songaila, A., Blades, J. C., Hu, E. M. & Cowie, L. L. *Astrophys. J.* **303**, 198-215 (1986).
12. Prévot, L. *et al. Astr. Astrophys. Suppl.* **62**, 23-37 (1985).

Quasars and superconducting cosmic strings

A. Vilenkin* & G. B. Field†

* Department of Physics and Astronomy, Tufts University, Medford, Massachusetts 02155, USA

† Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, Massachusetts 02138, USA

Loops of superconducting cosmic string acquire electrical currents in the magnetic fields of host galaxies and emit short bursts of highly directed electromagnetic radiation. Here, we propose that the jets observed in quasars are formed by particles accelerated to relativistic energies by such bursts, and that the central engines of quasars are therefore loops of cosmic string.

Cosmic strings are tubes of false vacuum formed at a phase transition in the early Universe¹. If the symmetry-breaking energy scale η is 10^{16} GeV , the mass per unit length of string is $\mu = 10^{22} \text{ g cm}^{-1}$, and $G\mu/c^2 = 10^{-6}$ (which determines the strength of gravitational interactions of strings). The gravitation of loops of string can form galaxies. In many grand unified theory (GUT) models strings are superconducting², and can support currents up to $i_{\text{max}} = e\eta/\hbar = 7 \times 10^{30} \text{ e.s.u. s}^{-1}$. Superconducting strings may explain phenomena in our Galaxy³ and, if primordial magnetic fields exist, the distribution of galaxies⁴.

Our quasar model is based on a loop of string located at the centre of the galaxy that formed around it. As quasars are present in greatest numbers at redshifts $z \sim 3$, corresponding to a cosmic

time $\sim 10^9 \text{ yr}$ in a universe of critical density⁵ (density parameter $\Omega = 1$), we concentrate on loops whose initial length (defined by M/μ where M is the mass of the loop) is $L_0 \sim 50 G\mu t/c \sim 50,000 \text{ light yr}$, which are the most numerous at that time (shorter loops having been destroyed by gravitational radiation¹). The 'radii' of such loops are $L_0/9 \sim 6,000 \text{ light yr}$ ⁶, and their masses and energies are $2.5 \times 10^{11} M_{\odot}$ and $5 \times 10^{65} \text{ erg}$, respectively.

The gas in galaxies is permeated by magnetic fields formed by dynamo action⁷. As the timescale for this is 10^8 - 10^9 yr , most galaxies are magnetized before the quasar era. As a string moves through a magnetic field it develops a current^{2,3} $i \sim \beta \xi \alpha v B T$, where β is a model-dependent constant of order unity, $\xi \sim 1$ is a correction for the effects of self inductance, α is the fine-structure constant, v (of order c) is the local velocity of the string, B is the magnetic field and T is a characteristic timescale. Initially, the length of the loop L_0 is greater than the length scale S ($\sim 3,000 \text{ light yr}$) over which galactic magnetic fields are expected to vary. As a result, i consists of harmonics of typical wavelength S . We therefore take $T \sim S/v$, and with $\beta = \xi = 1$, $B \sim 7 \times 10^{-6} \text{ G}$ (ref. 7) and $S \sim 3,000 \text{ light yr}$, the initial value of i is $i_0 \sim 5 \times 10^{24} \text{ e.s.u. s}^{-1}$. We define $j = i/i_{\text{max}}$, so $j_0 \sim 7 \times 10^{-7}$.

The loop shrinks as it loses energy by gravitational radiation according to $dL/dt = -50 G\mu/c$, and as it does so, the wavelengths of the harmonics of current decrease in proportion to L , and their amplitudes grow as L^{-1} , so that $j = j_0 L_0/L$. As this point is important for what follows, we argue briefly for it here. The current in a superconducting loop is proportional to the gradient of a massless scalar field φ defined on the two-dimensional world sheet of the loop^{2,8}. One can show from the conformal invariance of φ that its amplitude is not affected by

the shrinkage of the loop, while its wavelength is proportional to L . Hence the gradient of φ , and hence i , is proportional to L^{-1} .

As the loop shrinks, it will probably self-intersect and fragment into smaller loops. The efficiency of this process is not known; we shall assume that a substantial fraction of loops can shrink by a factor of ~ 100 without serious loss to fragmentation.

The radiation from a current-carrying string has recently been calculated (ref. 8 and D. Spergel, T. Piran and J. Goodman, Institute for Advanced Study preprint, December 1986). The mean power $P \sim 30\alpha\eta^2 j^{4/3}/\hbar = P_0(L/L_0)^{-4/3}$, where $P_0 = 3 \times 10^{44} \text{ erg s}^{-1}$, is dominated by emission of short bursts ($\Delta t \sim j^2 L/c$) of highly directed ($\theta \sim j^{2/3}$) energy from cusps formed on the string at which the velocity of the string momentarily reaches c . The bursts are emitted normal to the string and parallel to its motion; their energy is $PL/2c = 5 \times 10^{56}(L/L_0)^{-1/3} \text{ erg}$. The electric field E in such a burst satisfies $f \equiv eE\Delta t/mc \gg 1$ for protons at distances up to $10^7 L_0$. Under these conditions, test particles are accelerated to relativistic energies in the direction of the burst⁹; if the mean density ahead of a burst is n , a single burst accelerates to relativistic energies all particles up to a distance of $r = 5 \times 10^4(L/nL_0)^{1/3} \text{ light yr}$. With n of the order of 1 cm^{-3} , a few bursts will evacuate a cone of angle θ far beyond the galaxy. We identify such cones, repeatedly energized by bursts of low-frequency radiation, with the jets observed in quasars. Rees¹⁰ proposed a similar model for quasars but he assumed that the radiation is sinusoidal rather than in bursts, and attributed it to emission by rotating collapsed stars.

The observed radiation from jets is synchrotron emission from electrons in steady magnetic fields which have been drawn out of the host galaxy by the jet¹¹. In our model, the electrons are those accelerated to moderate energies in the outer parts of a burst (in which the field decreases as θ^{-3}). The parts of the loop not involved in cusps dissipate energy at a smaller rate. If the gas density is high, this dissipation will occur in relativistic shocks³. The energy of the shock-heated particles may escape through atomic radiative processes, or, if the latter fail to keep pace with the heating, the pressure will rise to the point that gas will start to escape from the galaxy. In the absence of detailed calculations, we assume that the former is the case.

As the loop shrinks, the emitted power increases according to $P_0(L/L_0)^{-4/3}$, reaching 3×10^{44} , 7×10^{45} , and $10^{47} \text{ erg s}^{-1}$ as L shrinks from L_0 to $0.1L_0$ to $0.01L_0$. The total energy radiated when the loop is longer than L is $3P_0t_0[(L/L_0)^{-1/3} - 1]$, where $t_0 = 10^9 \text{ yr}$; this reaches $1.3 \times 10^{62} \text{ erg}$ when $L \sim 10^{-2}L_0$. The time $t(P)$ over which the emitted power exceeds P is $t_0(P/P_0)^{-3/4}$; if the number of quasars is proportional to this time, the bolometric luminosity function is $\log \Phi = \text{constant} + 0.3M_{\text{BOL}}$, where M_{BOL} is the absolute bolometric magnitude.

At time t , loops of length $L > 50G\mu t/c$ have density $n_L = \Gamma(ct)^{-2}L^{-1} \text{ cm}^{-3}$, where Γ is a constant of the order of unity¹. In an $\Omega = 1$ universe, the total number of dying loops ($L \sim 50G\mu t/c$) with redshift z is $(54\pi\Gamma^2/50G\mu) \times \{z - 4[(1+z)^{1/2} - 1] + \ln(1+z)\}$.

We now compare the predictions of our model with observations. The total predicted energy is $\sim 10^{62} \text{ erg}$, compared with an estimate for the energy in particles and fields of $\sim 10^{61} \text{ ergs}$ for quasars and similar objects¹¹. Observed numbers of quasars increase with redshift and intrinsic faintness; evolutionary models¹² give $\sim 8 \times 10^6$ quasars with $M_B < -23$ (where M_B is the absolute blue magnitude) and $z < 3.5$, whereas we predict $2 \times 10^6 \Gamma$. The same models give a luminosity function at $z = 2.5$ in approximate agreement with our prediction if $M_B = \text{constant} + M_{\text{BOL}}$.

Radio components in quasars are observed¹³ with apparent superluminal velocities of $3c$ to $45c$. This phenomenon can be explained by relativistic motion with large γ if the jet is nearly along the line of sight¹⁴⁻¹⁶. We attribute the required γ to acceleration of electrons in the magnetic field enclosing the jet, by a burst with $v = c$. Extragalactic radio sources including quasars manifest jets that connect compact cores with more

extended lobe emission¹⁷. We attribute such jets to the cones of low-frequency radiation emitted by cusps on loops of string. As typical loops have two cusps¹⁸, jets should be accompanied by counterjets, as is observed for weak radio galaxies. In accordance with the usual interpretation^{16,19}, we attribute the absence of counterjets in quasars to Doppler dimming, as both the proximity to and the power of quasars imply large values of γ .

The fact that quasars are variable on timescales from seconds to years⁵ is consistent with our model. The duration of a burst is $\Delta t = j_0^2 L_0^2 / Lc \sim 1(L_0/L) \text{ s}$.

Because the loops we have been considering have very large initial masses, they are expected to promote the formation of massive galaxies and clusters. For quasars for which the underlying galaxy can be resolved, the galaxy is usually located in a group or cluster of galaxies of which it is the most luminous²⁰.

Our proposed model yields reasonable agreement with observations concerning the energy, number, luminosity, morphology, time variability, location of and motions in quasars. Although it is based on untested physical theory, it merits further study.

We thank Dr Wayne Christiansen for drawing our attention to the possible connection between radiation by strings and quasar jets, and Dr William Press, Dr Leonid Ozernoy and Dr Cedric Lacey for helpful comments. We acknowledge the support of NSF, NASA, and the General Electric Company.

Note added in proof: Dr Jeremiah Ostriker, Dr Martin Rees and Dr Craig Hogan have pointed out to us that our model would agree better with the observed numbers and sizes of quasars if $G\mu/c^2 = 10^{-7}$ rather than 10^{-6} .

Received 18 December 1986; accepted 6 February 1987.

- Vilenkin, A. *Phys. Rep.* **121**, 263-315 (1985).
- Witten, E. *Nucl. Phys.* **B249**, 557-592 (1985).
- Chudnovsky, E. M., Field, G. B., Spergel, D. N. & Vilenkin, A. *Phys. Rev.* **D34**, 944-950 (1986).
- Ostriker, J. P., Thompson, C. & Witten, E. *Phys. Lett.* **B180**, 231-239 (1986).
- Trimble, V. & Woltjer, L. *Science* **234**, 155-161 (1986).
- Turok, N. & Brandenburger, R. *Phys. Rev.* **D33**, 2175-2181 (1986).
- Sofue, Y., Fujimoto, M. & Wielebinski, R. *Astr. Astrophys.* **24**, 459-497 (1986).
- Vilenkin, A. & Vachaspati, T. *Phys. Rev. Lett.* **58**, 1041-1044 (1987).
- Ostriker, J. P. & Gunn, J. E. *Astrophys. J.* **157**, 1395-1417 (1969).
- Rees, M. J. *Nature* **229**, 312-317 (1971).
- Begelman, M. C., Blandford, R. D. & Rees, M. J. *Rev. mod. Phys.* **56**, 255-351 (1984).
- Schmidt, M. & Green, R. F. *Astrophys. J.* **269**, 352-374 (1985).
- Kellermann, K. I. & Pauliny-Toth, I. I. K. *Astr. Astrophys.* **19**, 373-410 (1981).
- Rees, M. J. *Mon. Not. R. astr. Soc.* **135**, 345-360 (1967).
- Ozernoy, L. M. & Sazanov, V. N. *Astrophys. Space Sci.* **3**, 395-423 (1969).
- Ginzburg, V. L. & Syrovatskii, S. I. *Astr. Astrophys.* **7**, 375-420 (1969).
- Bridle, A. H. & Perley, R. A. *Astr. Astrophys.* **22**, 319-358 (1984).
- Turok, N., *Nucl. Phys.* **B242**, 520-541 (1984).
- Ryle, M. & Longair, M. S. *Mon. Not. R. astr. Soc.* **136**, 123-140 (1967).
- Gehren, T., Fried, J., Wehinger, P. A. & Wyckoff, S. *Astrophys. J.* **278**, 11-27 (1984).

Possible large-scale structure at redshifts $\lesssim 0.5$

P. A. Shaver

European Southern Observatory, Karl-Schwarzschild-Strasse 2,
D-8046 Garching bei München, FRG

Large-scale structure has been measured using galaxies out to distances corresponding to redshifts of ~ 0.05 ¹. At much higher redshifts, marginal statistical evidence for the presence of clustering has been obtained using quasars^{2,3}, but the magnitude and evolution of this clustering are not yet well determined. In this paper evidence is given for the possible existence of very large-scale structure at intermediate distances, corresponding to redshifts $\lesssim 0.5$. This structure peaks in our direction of motion towards the microwave background, and could conceivably be the cause of it.

Structure on the largest scales is best studied using large, homogeneous quasar surveys covering many steradians of sky. Altogether the existing radio surveys cover almost the entire sky, but the surface density of objects varies strongly from one survey to another, depending on the limiting sensitivity. However, the redshift distribution of radio quasars is not such

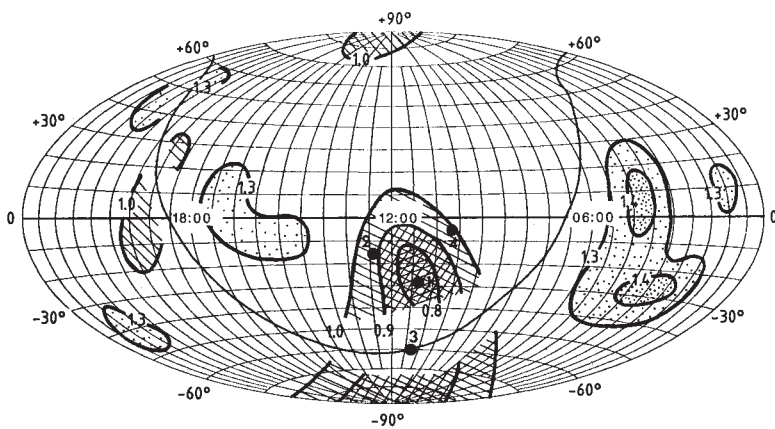


Fig. 1 Sky distribution in celestial coordinates of median redshift for all quasars in the Véron catalogue with 6-cm flux density > 0.1 Jy and confirmed redshifts from slit spectroscopy. The bin size is 0.5 steradian, containing typically 30–40 quasars. Hatched and dotted regions have median redshifts < 1.0 and > 1.3 respectively. The galactic plane is indicated, and forms a 'zone of avoidance' due to obscuration. The direction of motion towards the microwave background¹ (1), two estimates of the direction of motion of galaxies out to $6,000 \text{ km s}^{-1}$ (refs 6–8) (2, 3), and the location of the peak in the distribution of IRAS galaxies⁵ (4) are also marked.

a strong function of survey sensitivity, and so may give some indication of possible large-scale structure.

Figure 1 shows the distribution of median redshift for all quasars with confirmed redshifts and 6-cm flux density > 0.1 Jy in a current version of the Véron catalogue of quasars⁴ (a flux density limit is used simply to maximize the probability that the quasars were discovered first as radio sources, and thereby to minimize optical selection effects). Given the heterogeneous nature of the sample, coming from many different radio surveys, it is impressive how flat the distribution is over the sky. There is no significant correlation between median redshift and median flux density. The all-sky median redshift is 1.15, and the deviations from this value do not exceed 0.15 over most of the sky. Even at galactic latitudes less than 15° , where the greater difficulty of identifications might be expected to have some effect, the median redshift is still the same. The most striking deviation, which has a probability of $\leq 10^{-3}$ of being a statistical fluctuation (from both Monte Carlo and Mann-Whitney tests), occurs near 11 h–30°, and it is immediately intriguing that this coincides with our direction of motion towards the microwave background within $\sim 5^\circ$. Thus, although this map should not be used rigorously because of possible selection effects, it does suggest that further investigation using more homogeneous samples may be worthwhile.

Fortunately the regions of greatest interest are at declinations in the range 0° to -40° , which are covered by the deep and extensive Parkes⁹ and Molonglo¹⁰ surveys. These radio surveys are complete over large areas, although the identifications of the radio sources are far from complete. Major efforts have been made to identify radio sources from the Parkes 11-cm survey, so it is reasonable to focus on these: in particular the zones 22–05 h and 10–15 h in the declination range -4° to -45° , which have been completely surveyed down to 0.25 Jy at 11 cm. The surface density of known quasars with 11-cm flux density > 0.25 Jy is virtually identical in the two zones. In the 22–05 h zone there are 13 such quasars with $z < 0.6$, and 86 with $z > 0.6$. Thus, in the 10–15 h zone we would expect 9 such quasars with

$z < 0.6$, and 61 with $z > 0.6$. Instead we find 24 and 43 respectively.

The zone 22–05 h includes the south galactic pole. This region is particularly favourable for identification work, as it has a relatively low stellar density and extinction. It has been intensively studied. The 10–15 h zone, on the other hand, is at relatively low galactic latitude; identifications there are likely to be less complete, particularly for high-redshift quasars which are on average relatively faint. Thus, while a relative deficiency of high-redshift quasars might be expected in the 10–15 h zone, an excess of the relatively bright low-redshift quasars is difficult to explain as a selection effect. This excess can be clearly seen in Fig. 2.

If there really is an excess of low-redshift radio quasars in the 10–15 h zone, then there should also be an excess of radio sources generally in this direction. It would, of course, be greatly diluted because of radio sources at other redshifts, and may not be of high significance, but it would at least be free of the selection effects involved in the optical identifications. This prediction can easily be tested by reference to the original Parkes 11-cm survey⁹. In the 10–15 h zone above the galactic plane the surface density of sources with 11-cm flux density > 0.25 Jy is 11% higher than elsewhere in the regions of completeness; from a χ^2 test the probability that this is a statistical fluctuation is < 0.02 .

Similarly the Molonglo catalogue¹⁰, which is complete above 1 Jy at 408 MHz from -85° to $+18.5^\circ$ (excluding the region within 3° of the galactic plane), can be used, although being a low-frequency survey it is not as sensitive to flat-spectrum quasars, so any excess may be further diluted. The surface density of sources in the 10–15 h zone ($-45^\circ < \delta < -4^\circ$) is 7% higher than the average elsewhere, and the probability that this is a statistical fluctuation is < 0.05 . Thus the prediction of an excess of radio sources in the 10–15 h zone appears to have been confirmed. Although part of this excess could be due to relatively local radio sources associated with concentrations of visible galaxies, it is at least consistent with the suggestion that

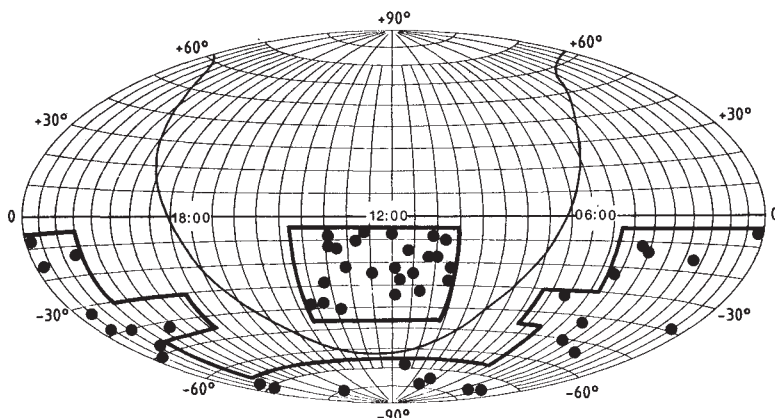


Fig. 2 Distribution of quasars with redshifts < 0.6 and 11-cm flux density > 0.25 Jy over the region (enclosed by heavy lines) in which the Parkes catalogue is complete to that limit of flux density.

there is a real excess of low-redshift quasars in that region.

Further study of this excess will require a variety of new data. A complete sample of radio quasars covering the relevant areas, with complete identifications and confirmed redshifts, would help to dispel any lingering doubts regarding selection effects. Deep ultraviolet-excess or objective-prism quasar surveys in the 10–15 h zone, in conjunction with others made under identical conditions in other areas of the sky, would show any differences in redshift distribution at a high level of significance. [Unfortunately there are very few optically selected (non-radio) quasars in the 10–15 h zone; the two studies^{11,12} in which more than a few quasars have been confirmed were both based on objective-prism data and geared specifically towards high-redshift (emission-line) objects, and are not useful in examining any possible excess of low-redshift quasars.] Source counts from deep radio and X-ray surveys would also help in studying differences between the populations in this direction and others. The X-ray background may also provide a constraint, although its origin is still a matter of debate; a weak (0.5%) dipole appears to be present, again coincident with the direction of motion towards the microwave background¹³.

The structure implied by the distribution of low-redshift quasars may be complex—there is some indication of a relative deficiency in the direction of the south galactic pole—but it appears to be dominated by the excess in the 10–15 h zone. This feature has an overdensity of ~ 1 –2, angular diameter $\sim 30^\circ$ – 40° , and mean redshift ~ 0.3 – 0.4 , although it is difficult to determine these parameters with any precision. At the corresponding distance of $\sim 800 h^{-1}$ Mpc it would have a diameter of $\sim 400 h^{-1}$ Mpc.

It may be significant that this large feature lies in the direction of our motion towards the microwave background (Fig. 1). It has recently been suggested^{6–8} that distant galaxies out to $\geq 6,000 \text{ km s}^{-1}$ share this motion. If indeed there is a bulk stream-

ing motion in a region over $100 h^{-1}$ Mpc diameter surrounding us, the obvious interpretation is that it is due to the attraction of a very large mass lying beyond, in the direction of motion¹⁴. While asymmetries in the streaming motion suggest that the mass may not be too far beyond^{7,8}, the far more distant structure implied by the excess of low-redshift quasars (assuming they follow the mass) would also suffice to produce the 600 km s^{-1} streaming motion. From linear perturbation theory¹⁵, this velocity can be produced with the parameters (overdensity, angular size and redshift) given above and a value for the cosmological density parameter Ω of the order of unity. Such an association between a distant mass and local kinematics would provide new information on large-scale properties of the universe.

I thank Ann Savage, Bob Sanders, Jasper Wall, and many other colleagues for helpful discussions. I am grateful for the hospitality of the CSIRO Division of Radiophysics and the Raman Research Institute, where part of this work was carried out.

Received 18 December 1986; accepted 24 February 1987.

1. Oort, J. H. A. *Rev. Astr. Astrophys.* **21**, 373–428 (1983).
2. Shaver, P. A. *Astr. Astrophys.* **136**, L9–L10 (1984).
3. Shanks, T., Fong, R., Boyle, B. J. & Peterson, B. A. in *Quasars* (eds Swarup, G. & Kapahi, V. K.) 37–43 (Reidel, Dordrecht, 1986).
4. Véron-Cetty, M.-P. & Véron, P. ESO Scientific Report No. 4 (1985).
5. Yahil, A., Walker, D. & Rowan-Robinson, M. *Astrophys. J.* **301**, L1–L5 (1986).
6. Collins, C. A., Joseph, R. D. & Robertson, N. A. *Nature* **320**, 506–508 (1986).
7. Burstein, D. et al. in *Galaxy Distances and Deviations from Universal Expansion* (eds Madore, B. F. & Tully, R. B.) 123–130 (Reidel, Dordrecht, 1986).
8. Dressler, A. et al. *Astrophys. J.* **313**, L37–L42 (1987).
9. Bolton, J. G., Savage, A. & Wright, A. E. *Aust. J. Phys. Astrophys. Suppl.* **46**, 1–27 (1979).
10. Large, M. I., Mills, B. Y., Little, A. G., Crawford, D. F. & Sutton, J. M. *Mon. Not. R. astr. Soc.* **194**, 693–704 (1981).
11. Osmer, P. S. & Smith, M. G. *Astrophys. J.* **213**, 607–618 (1977).
12. Kunth, D., Sargent, W. L. W. & Kowal, C. *Astr. Astrophys. Suppl.* **44**, 229–239 (1981).
13. Shafer, R. A. & Fabian, A. C. in *Early Evolution of the Universe and its Present Structure* (eds Abell, G. O. & Chincarini, G.) 333–343 (Reidel, Dordrecht, 1983).
14. Lynden-Bell, D. *Q. Jl R. astr. Soc.* **27**, 319–327 (1986).
15. Davis, M. & Peebles, P. J. E. *A. Rev. Astr. Astrophys.* **21**, 109–130 (1983).

Magnetospheric charge-exchange effect on the electroglow of Uranus

W. H. Ip

Max-Planck-Institut für Aeronomie, D-3411 Katlenburg-Lindau, FRG

One common feature of the magnetospheric systems of the outer planets explored by the Pioneer 10 and 11 and the Voyager 1 and 2 spacecraft is that extended distributions of neutral gas of planetary exospheric origin and/or of satellites—rings origin coexist with the magnetospheric plasma. This property invariably leads to the occurrence of charge-exchange loss of the charged particles in the inner magnetospheres of Jupiter, Saturn and Uranus. With emphasis on the Uranian system I examine here the idea that the injection of the recombined fast neutrals into the planetary upper atmosphere might be contributing in part to the generation of the dayside electroglow. The recent results from the Voyager 2 encounter also permit the idea that the associated impact ionization effect could be important in maintaining the nightside ionosphere.

The Voyager 2 encounter with Uranus¹ has produced a wealth of new information on this gaseous outer planet, its satellite and ring systems and a new type of magnetosphere. Because of the unique pointing direction of Uranus and the large tilt angle ($\theta \sim 60^\circ$) between the dipole moment and the rotational axis², the general pattern of magnetospheric convection is now believed to be similar to that of the terrestrial magnetosphere but with the whole system in co-rotation with the planet^{3–6}. In other words, unlike the cases with the jovian and saturnian magnetospheres^{7,8}, the sunward convection effect driven by a dawn-to-dusk electric field can penetrate very deep into the magnetosphere.

At the same time, according to the plasma science experiment³ and the planetary radioastronomy experiment⁹, there might be indications for the formation of a plasmopause separating the sunward-convecting plasma from the 'co-rotating' plasmasphere. Thus the streamlines (or equipotential lines) do not necessarily follow a rectilinear pattern with some of them directly intercepting the planet on the nightside hemisphere—as expected in the idealized situation with the tilt angle, $\theta \sim 90^\circ$. The presence of a plasmopause structure at Uranus, however, is a complicated issue still to be investigated in detail.

Another result of major importance to atmospheric and magnetospheric physics has to do with the ultraviolet spectroscopy (UVS) observations of the so-called electroglow from the dayside hemisphere of Uranus^{10,11}. The intense ultraviolet emissions from the H₂ and H I Rydberg series were found to be uniformly distributed over the dayside disk and the H Lyman- α emission brightness of 1.5 KR near the subsolar point is consistent with the International Ultraviolet Explorer (IUE) time-averaged value of 1.4 KR¹². The characteristic property of non-solar excitation of the extreme ultraviolet (EUV) emissions (as found at Jupiter and Saturn) is also established in the case of Uranus¹¹. One interesting difference, however, is that the H₂ emission spectra of Jupiter and Saturn can be interpreted as being the result of impact excitation by electrons of about 30 eV energies whereas that of Uranus can be fitted by invoking electrons with energies ~ 3 eV. The total power input required for the excitation of the electroglows (1.5×10^{12} W for Saturn and 1.7×10^{11} W for Uranus) might explain the elevated temperatures of the exospheres (~ 750 K for both Saturn and Uranus).

One new piece of information which might be essential to the understanding of the generation mechanisms of the electroglows concerns the fact that the altitude distributions of the ultraviolet emissions from limb-scanning measurements suggest that part of the excitations could be related to the collisional impact effect

of energetic protons and/or fast hydrogen atoms impinging on the upper atmosphere of Saturn and Uranus^{11,13}.

Because of the ring absorption effect, energetic charged particles in the saturnian magnetosphere cannot diffuse inwards and reach the planetary atmosphere without undergoing nearly complete depletion¹⁴; the only effective way to inject energetic protons or fast hydrogen atoms then is via the process of charge-exchange recombination in the extended neutral gas cloud surrounding the planet^{15,16}. In view of the extensive structure of the hydrogen corona at Uranus as revealed by the Voyager EUV observations¹¹, charge-exchange processes could be equally important in contributing to the electroglow emission. In the following, the physical model as well as a first-order estimate of the energy disposition rate to the upper atmosphere of Uranus will be described to evaluate the validity of this idea.

To begin with, note that the low-energy charged particle (LECP) experiment has determined that the magnetosphere of Uranus may be characterized as a low β plasma⁴. The maximum value of the ratio of the charged particle kinetic energy density (ϵ) to the magnetic field energy density ($B^2/8\pi$) as measured by Voyager 2 is $\beta \leq 0.1$ at $L \sim 15$. Taking this as a limit, we have as an extrapolation to the ring current particle flux in regions inside of the orbit of Miranda not scanned by the Voyager 2 spacecraft

$$\epsilon(L) \sim \beta B_0^2/8\pi L^6 \quad (1)$$

where $\beta < 0.1$ and B_0 ($=0.23$ gauss) is the equatorial surface field. With this expression, the total charged particle energy in a volume between L_a and L_b can be written as

$$E_{\text{total}} \sim \delta \int_{L_a}^{L_b} \epsilon(L) dV \sim \delta R_U^3 \int_{L_a}^{L_b} \epsilon(L) \cdot 2\pi L^2 dL \quad (2)$$

where δ is a form factor adjusting for the detailed distribution of the charged particles in the magnetic field. For $L_a \sim 2$ and $L_b \sim 5$, $E_{\text{total}} \sim 0.1 \beta \times \delta \times B_0^2 \times R_U^3$ or $\beta \times \delta \times 10^{26}$ erg.

The limits of $L_a \sim 2$ and $L_b \sim 5$ are used in the above calculation as the charge-exchange process should be most efficient in this radial interval because of the large concentration of the neutral hydrogen atoms therein. Although the analytical expression of n_H as given in Broadfoot *et al.*¹¹ for L between 1.6 and 2 would predict a very small hydrogen density at $3-5 R_U$ ($n_H < 1 \text{ cm}^{-3}$), Shemansky and Smith¹⁰ derived density curves (with $n_H \sim 10-100 \text{ cm}^{-3}$ for L between 2 and 4) corresponding to the hydrogen atoms ejected from the exobase in association with electroglow excitation. Using $n_H \sim 10 \text{ cm}^{-3}$ as a nominal value, the timescale for charge-exchange recombination is

$$t_{\text{cx}} \sim (n_H \sigma V)^{-1} \sim 10^6 \text{ s} \quad (3)$$

where $\sigma V \sim 10^{-7} \text{ cm}^3 \text{ s}^{-1}$ at particle energies $\sim 15 \text{ keV}$ (ref. 17). The energy dissipation rate of the ring current is thus (see ref. 17 for a review)

$$\dot{E}_{\text{cx}} \sim \frac{E_{\text{total}}}{t_{\text{cx}}} \sim \left(\frac{\beta}{0.1}\right) \times \left(\frac{\delta}{1}\right) \times \left(\frac{n_H}{10}\right) \times 10^{13} \text{ W} \quad (4)$$

Numerical calculations employing the Monte Carlo method similar to those by Prag *et al.*¹⁸ and Prölss¹⁹ have been performed for a centred dipole field and the results indicate that about 3-5% of the recombined energetic neutral particles will impact the planetary surface with concentration near the equator. The energy disposition rate to the dayside upper atmosphere is thus

$$\dot{E}_{\text{cx}}(\text{surface}) \sim \left(\frac{\beta}{0.1}\right) \times \left(\frac{\delta}{1}\right) \times \left(\frac{n_H}{10}\right) \times 5 \times 10^{11} \text{ W} \quad (5)$$

In the optimal condition with $\beta \sim 0.1$, $\delta \sim 1$, and $n_H \sim 10 \text{ cm}^{-3}$, $\dot{E}_{\text{cx}}(\text{surface}) \sim 5 \times 10^{10} \text{ W}$. In comparison, the energy inputs for the electroglow and nightside aurora are about 10^{11} W each at Uranus¹¹. If the impact flux is assumed to be uniform over the disk, the energy flux contributed by the recombined fast neutrals should be of the order of $0.02 \text{ erg cm}^{-2} \text{ s}^{-1}$ at maximum. This

means that the fast neutrals from the charge exchange loss of the ring current system perhaps could be a significant source of excitation of the electroglow at Uranus which requires a total disposition rate of about $0.08 \text{ erg cm}^{-2} \text{ s}^{-1}$.

Because straggling of the fast neutrals in the upper atmosphere would be followed by a succession of charge-exchange interaction, excitation (and hence atmospheric heating) and ionization of the neutral molecules via electron ejection²⁰, one consequence of charge-exchange processes in the magnetosphere would be an additional ionization of the H_2 atmosphere of Uranus. On the dayside hemisphere, this contribution may be easily masked by other effects; however, on the nightside hemisphere which is blocked from sunlight, the injection of fast neutrals could be partially responsible for the maintenance of a nightside ionosphere as revealed by the detection of Uranus electrostatic discharges (UED) by the planetary radio astronomy experiment²¹. The required energy disposition rate should be very much smaller than $0.02 \text{ erg cm}^{-2} \text{ s}^{-1}$.

In the above discussion with emphasis on the possible relation between the fast neutral injection effect and the electroglows on Saturn and Uranus, limitations on observational data have rendered a number of uncertainties in the numerical estimates. For example, the adopted β value (~ 0.1) could have been too large and the assumed density distribution of the neutral hydrogen atoms in the extended corona may have been atypical. (Note that $\beta \sim 0.01$ and $t_{\text{cx}} \sim 10^5 \text{ s}$ would yield the same value of $\dot{E}_{\text{cx}}$ (surface).) On the other hand, some of the parameters used might be more appropriate for a disturbed interval of the magnetosphere rather than the quiet time situation. Furthermore, no account has been taken of the possibility that radial transport of the energetic charged particles could continue until reaching the upper atmosphere such that the efficiency of surface impact would be significantly enhanced. Also, the role of magnetospheric electrons is not discussed here, even though they could be important in maintaining a certain level of atmospheric ionization during the night-time²². There are therefore a number of open issues in need of detailed investigations. A very important factor as discussed here is the UVS observations of the signature of energetic protons-fast hydrogen impact excitations at Saturn and Uranus^{11,15}. In a certain way, they could be regarded as remote sensing the charge-exchange process in the ring current systems. A determination of the corresponding energy disposition rates would clarify to what extent the energetics of the ring currents are coupled to the upper atmospheric dynamics of the outer planets.

Finally, note that additional energy sources other than charge-exchange loss of the ring current alone should be involved to account for the strong day-night asymmetry. Indeed, there is no obvious way that the charge-exchange process could be modulated so that it takes place predominantly only in the dayside hemisphere. The present estimate nevertheless suggests that this particular coupling effect between the magnetospheric plasma and the extended neutral exosphere could be quite significant in the excitation of atmospheric emission as well as heating and ionization of the nightside disk of Uranus.

Received 7 November 1986; accepted 20 January 1987.

1. Stone, E. C. & Miner, E. D. *Science* **233**, 39-43 (1986).
2. Ness, N. F. *et al. Science* **233**, 85-89 (1986).
3. Bridge, H. S. *et al. Science* **233**, 89-93 (1986).
4. Krimigis, S. M. *et al. Science* **233**, 97-102 (1986).
5. Vasyliunas, V. M. *Geophys. Res. Lett.* **13**, 621-623 (1986).
6. Selesnick, R. S. & Richardson, J. D. *Geophys. Res. Lett.* **13**, 624-627 (1986).
7. Brice, N. M. & Ioannidis, G. A. *Icarus* **13**, 173-183 (1970).
8. Brice, N. M. *J. geophys. Res.* **72**, 5193-5211 (1967).
9. Warwick, J. W. *et al. Science* **233**, 102-106 (1986).
10. Shemansky, D. E. & Smith, G. R. *Geophys. Res. Lett.* **13**, 2 (1986).
11. Broadfoot, A. L. *et al. Science* **233**, 74-79 (1986).
12. Clarke, J. T. *et al. J. geophys. Res.* (in the press).
13. Yelle, R. V., Sandel, B. R., Shemansky, D. E. & Kumar, S. *J. geophys. Res.* **91**, 8756-8770 (1986).
14. Fillius, W., Ip, W.-H. & McIlwain, C. E. *Science* **207**, 425-431 (1980).
15. Broadfoot, A. L. *et al. Science* **212**, 206-211 (1981).
16. Ip, W.-H. in *Planetary Rings* (ed. Brahic, A.) 575-596 (Cepadues-Editions, Toulouse, 1984).
17. Tinsley, B. A. *J. atmos. terr. Phys.* **42**, 227 (1981).

18. Prag, A. B., Morse, F. A. & McNeal, R. J. *J. geophys. Res.* **71**, 3141-3154 (1966).
 19. Pröls, G. W. *Ann. Geophys.* **29**, 503-508 (1973).
 20. Rees, M. H. *Planet. Space Sci.* **30**, 463-472 (1982).
 21. Zarka, P. & Pedersen, B. M. *Nature* **323**, 605-608 (1986).
 22. Lyons, L. R. & Richmond, A. D. *J. geophys. Res.* **83**, 2201-2204 (1978).

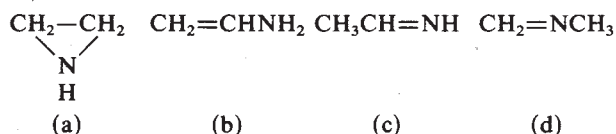
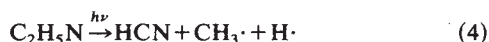
HCN and chromophore formation on Jupiter

James P. Ferris & Yoji Ishikawa

Department of Chemistry, Rensselaer Polytechnic Institute, Troy, New York 12180-3590, USA

The formation of HCN¹ and chromophores² are two of the major unsolved problems of the atmospheric chemistry of Jupiter. The question to be dealt with is the same in each case: how can these unsaturated organic compounds be formed in the highly reducing atmosphere² (89% H₂) present on Jupiter? The photolysis of ammonia/acetylene mixtures provides an answer to this question. Here we report the formation of both HCN and chromophores along with experimental data which support the premise that this photochemical process provides a route for the formation of both substances. It is not clear whether significant amounts of HCN are also formed by lightning on Jupiter^{3,4}.

It was first suggested by Kaye and Strobel⁵ that HCN is formed on Jupiter by the photochemical addition of ammonia to acetylene. Reaction sequence (1)–(4) was outlined and compounds (a–d) were proposed as possible structures for C₂H₅N, the proposed HCN precursor formed in reaction (3). Neither HCN nor compounds (a–d) have been reported as products in previous investigations of the photolysis of ammonia/acetylene mixtures^{6,7}.



HCN is a product of the photolysis of ammonia in the presence of acetylene at room temperature (Table 1). Initial experiments (not described) established that HCN was formed from ammonia/acetylene mixtures where ammonia absorbed more than 98% of the light from a 185 nm source and more than 99.9% of the light from a 206.2 nm source (ref. 8). In addition, a brown, presumably oligomeric, deposit forms on the photolysis cell window.

Table 1 Variations in the yield of HCN, brown oligomers and HCN precursors

Temperature (K)	H ₂ (torr)	HCN* (m × 10 ⁵)	Brown oligomers†	HCN precursors‡
298	0	10.1 (0.9)	0.24 (0.005)	0
298	77.5	6.4 (2.7)	0.09 (0.04)	0
178	0	2.9 (0.75)	0.14 (0.01)	0.12 (0.03)
178	77.5	0.55 (0.13)	0.09 (0.04)	0.09 (0.12)
178	700	0.39 (0.14)	0.05 (0.01)	0.27 (0.10)

In these experiments NH₃ at 5 torr and 2.5 torr acetylene were irradiated for 8 h; standard deviations given in parentheses.

* Monitored by the method of Kruse and Mellon^{9,10}.

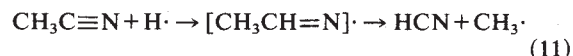
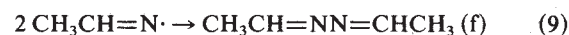
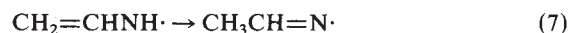
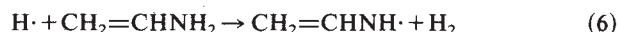
† Monitored by absorbance at 300 nm.

‡ Monitored by absorbance at 220 nm.

Photolyses were performed in the presence of hydrogen and at lower temperatures to determine if the same products were formed using conditions close to those prevalent in the atmosphere of Jupiter (Table 1). The HCN yield decreased by 70% when the temperature was lowered to 178K and there was a further 80% decrease in the yield when 77.5 torr of hydrogen was added to the photolysis mixture. No further decrease was observed when 700 torr of hydrogen was used instead of 77.5 torr (Table 1).

A HCN precursor fraction was detected by an enhanced ultraviolet absorption at 220 nm when the photolyses were performed at 178K and the photolysis cell was warmed to room temperature for UV analysis (Table 1). These compounds condense on the cell wall at 178K where they are no longer subject to ultraviolet irradiation. Combination gas chromatographic-mass spectral analysis established that acetonitrile (e) and acetaldehyde ethylidenehydrazone (f) were the major nitrogen-containing products in the precursor fraction. Compounds (a–d) were not detected in this fraction.

A reaction pathway to HCN can be proposed on the basis of this and other research described below^{11,12}. The reaction process is initiated by the photodissociation of ammonia (1) since it was established that it is absorbing more than 98% of the light. The addition of a hydrogen atom to acetylene (2) is the next step since the pseudo first order rate constant for hydrogen atom addition¹¹ is 3,000 times greater than NH₂· addition at 178K¹². Steps (5)–(9) provide a concise and plausible speculative rationale for the formation of the observed HCN precursors (e) and (f).



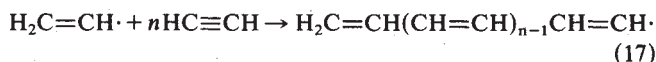
The photolysis of both precursors ultimately results in the formation of HCN. Irradiation of (f) results in the formation of acetonitrile (e) with a quantum yield of 0.5 (10) (see ref. 13). The hydrogen-atom-initiated conversion of acetonitrile to HCN (11) proceeds with high efficiency and does not require light absorption by the acetonitrile¹⁴. We observed that HCN is formed by the photolysis of 5 torr ammonia in the presence of 0.3 torr of acetonitrile using a 185 nm light source.

The vapour pressure of acetonitrile was estimated to ascertain that its photolysis to HCN is a plausible reaction in the Jovian atmosphere. Since there are no reports of the vapour pressure of acetonitrile in the 150–200K range, its vapour pressure was estimated by comparison with that of the vapour pressure of HCN. The vapour pressure of HCN was calculated¹⁵ to be 4 × 10^{−4} torr at 150K which is equivalent to a mixing ratio of 5 × 10^{−7} at 1 bar. This vapour pressure is 250 times greater than the measured mixing ratio on Jupiter¹ of 2 × 10^{−9}. Since the vapour pressure of acetonitrile is one-tenth that of HCN in the 220–273K range¹⁶, it can be estimated that the vapour pressure of acetonitrile at 150K is one-tenth that of HCN or 25 times greater than the observed HCN mixing ratio. It can be concluded that acetonitrile is sufficiently volatile at 150K for the photochemical generation of the levels of HCN observed in the atmosphere of Jupiter.

The yield of chromophoric oligomers decreases by about 60% when 77.5 torr of hydrogen is added to the photolysis mixture (Table 1). Little or no additional decrease in oligomer formation is observed on addition of 700 torr hydrogen or lowering the temperature to 178K. These findings suggest that the hydrogen-

atom-initiated oligomerization of acetylene is a source of chromophores on Jupiter.

It was proposed in 1979 that the photo-oligomerization of acetylene is a source of chromophores on Jupiter¹⁷. This explanation is no longer valid in view of the more recent data by Okabe¹⁸ where he studied the effect of pressure and hydrogen on the photochemical conversion of acetylene to diacetylene. Dissociation of the singlet excited state of acetylene yields the ethynyl radical (12) which adds to the ground state acetylene to give diacetylene (13). The reaction of the metastable acetylene (vinylidene)¹⁹ formed in (14) with ground state acetylene (15) only proceeds at an acetylene partial pressure greater than 0.7 torr because the vinylidene is quenched to acetylene by other gases faster than it adds to acetylene. The reaction of the ethynyl radical formed in (12) with acetylene (13) is a potential route to oligomers. The rate of reaction (13) is 7×10^3 times faster than the rate of reaction of the ethynyl radical with hydrogen (16) at 170K¹⁸. But, as the hydrogen acetylene ratio² on Jupiter is 9×10^6 , reaction (16) will predominate and the oligomerization of acetylene will not proceed efficiently in laboratory studies¹⁷ and on Jupiter.



It can also be concluded that the hydrogen-atom-initiated oligomerization of acetylene will proceed at the high partial pressure of hydrogen on Jupiter by consideration of the mechanism of the condensation reactions. The first steps, (1) and (2), are the same as those proposed for the formation of HCN. Oligomerization then proceeds in a stepwise fashion via vinyl radicals (17). The vinyl radical intermediates react with molecular hydrogen (18) 10^8 times more slowly²⁰ than do the ethynyl radicals (16) formed by the direct photolysis of acetylene¹⁸. Thus the hydrogen atom-initiated oligomerization of acetylene proceeds in the presence of a high partial pressure of hydrogen in laboratory experiments (Table 1) and on Jupiter.

This work was supported by NASA grant NAGW476.

Received 11 November 1986; accepted 19 January 1987.

1. Tokunaga, A. T., Beck, S. C., Geballe, T. R., Lacy, J. H. & Serabyn, E. *Icarus* **48**, 239-289 (1981).
2. Strobel, D. F. *Int. Rev. Phys. Chem.* **3**, 145-176 (1983).
3. Bar-Nun, A. & Podolak, M. *Icarus* **64**, 112-124 (1985).
4. Lewis, J. S. & Fegley, M. B. Jr *Space Sci. Rev.* **39**, 163-192 (1984).
5. Kaye, J. A. & Strobel, D. F. *Icarus* **54**, 417-433 (1983).
6. Tsukada, M., Oka, T. & Shida, S. *Chem. Lett.*, 437-440 (1972).
7. Masanet, J., Vermeil, C. & Amsterdamski, C. Pap. presented at Meet. of le Comité de L'ATP de Planétologie L'Institut National D'Astronomie et de Géophysique, Paris, February 1982, 69 (Abstr.).
8. Harteck, P., Reeves, R. R. & Thompson, B. A. *Z. Naturf.* **19A**, 2-6 (1984).
9. Kruse, J. M. & Mellon, M. G. *Analyt. Chem.* **25**, 446-450 (1953).
10. Chen, C. T. thesis, Boston Univ. (1973).
11. Payne, W. A. & Stief, L. J. *J. chem. Phys.* **64**, 1150-1155 (1976).
12. Bosco, S. R., Navo, D. F., Brobst, W. D. & Stief, L. J. *J. chem. Phys.* **81**, 3505-3511 (1984).
13. Brinton, R. K. *J. Am. chem. Soc.* **77**, 842-846 (1955).
14. McElcheran, D. E., Wijnen, M. H. J. & Steacie, E. W. R. *Can. J. Chem.* **36**, 321-329 (1958).
15. Appleton, G. T. & Van Hook, W. A. *J. chem. Engng Data* **27**, 363-365 (1982).
16. Stull, D. P. *Ind. Engng Che.*, **39**, 517-540 (1947).
17. Noy, N., Bar-Nun, A. & Podolak, M. *Icarus* **40**, 199-204 (1979).
18. Okabe, H. *J. chem. Phys.* **76**, 2772-2778 (1981); *Can. J. Chem.* **61**, 850-855 (1983).
19. Laufer, A. H. *J. Photochem.* **27**, 267-271 (1984).
20. Yung, Y. L. & Strobel, D. F. *Astrophys. J.* **239**, 395-402 (1980).

Speckle patterns permit direct observation of phase breaking

Moshe Kaveh, Michael Rosenbluh & Isaac Freund

Department of Physics, Bar-Ilan University, Ramat-Gan, Israel

Anderson¹ and Mott² showed that scattering from a strongly disordered system qualitatively changes the way in which electron waves propagate. It has recently been realized that these changes occur universally for all waves, in particular optical waves³⁻⁵. It is also becoming apparent that many important new phenomena previously inferred indirectly for electrons can now be observed directly for photons. Here, we show how the time evolution of the scattered wave due to inelastic phase breaking (dephasing) processes can be followed directly. When coherent light is scattered by a disordered system an image of the emergent wave is obtained in the form of an apparently random intensity alternation known as speckle. From the time evolution of this speckle pattern we have found that multiple scattering greatly reduces the time required for phase breaking. This unexpected result is explained in terms of the extreme fragility of the phases of long trajectories, and appears to have important implications for a wide variety of phenomena.

Multiple elastic scattering of electron or optical waves in a random system produces strong interference which, under suitable conditions, can lead to a stable, localized wave packet⁶⁻¹⁰. As a result of inelastic scattering, the phase coherence needed for the constructive interference that produces localization is degraded in a time τ_ϕ known as the phase-breaking time^{11,12}. Following its destruction, the wave packet is, of course, reconstructed in a different form and place owing to subsequent elastic scattering events, so that the nature of the wave function¹³

in a random medium can be strongly influenced by phase breaking. The process of multiple scattering may be described as a sequence of single scattering events, each with an average inelastic energy transfer (gain or loss) of ΔE . It is reasonable to suppose that after N scattering events, which could be described as an N -step random walk in energy space, the net energy transfer would be of order $N^{1/2}\Delta E$. The uncertainty principle relates time and energy through $\Delta E\tau \sim \hbar$, so one might anticipate that the phase-breaking time for an N -fold multiple-scattering process, where $\tau_\phi = \tau_m$, would be shorter than the phase-breaking time for single scattering, where $\tau_\phi = \tau_1$, by this same factor of $N^{1/2}$. Although this analysis appears to be almost self-evident, we show later that a more careful treatment carried out directly in the time domain leads to the important conclusion that the reduction in phase-breaking time is proportional to N itself, and not to $N^{1/2}$. We find from an analysis of our experimental data that the appropriate value of N for photon backscattering is $N \approx 25$.

Although the process of phase breaking is extremely difficult to observe directly for electrons⁸⁻¹⁰ (it is usually inferred indirectly from the temperature dependence of the conductivity), it is easily studied for photons through the time evolution of speckle patterns, as such patterns image the spatial variation of the emergent wave. This new method permits direct visual observation of the formation of a given wave configuration, its decay in time, and the subsequent formation of a new, uncorrelated configuration. The phase-breaking time is just the time over which the decay of the wave occurs. We studied this fascinating behaviour by observing coherent light scattered from randomly diffusing particles in which essentially all scattering events are weakly inelastic due to a small Doppler shift that slightly changes the energies of the scattered photons. For a very dilute system of scatterers in which only single scattering is important, the characteristic phase-breaking time τ_1 may be shown to be the average time required for a typical scatterer to

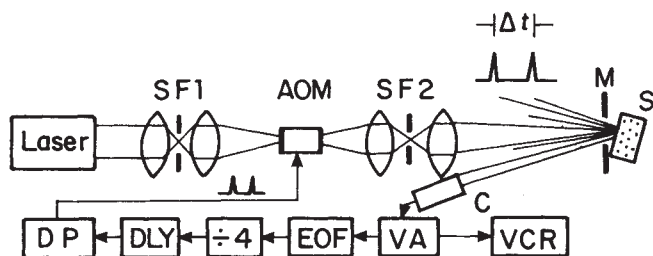


Fig. 1 Optical system and video-recording apparatus. The laser is a 2-W Ar-ion laser operated at a wavelength of 514.5 nm. SF1 and SF2 are spatial filters, AOM is a high-contrast, acousto-optic modulator used to produce short optical-pulse pairs of time separation Δt , and M is a mask which eliminates effects due to total internal reflection in the window of sample cell S. The video-recording apparatus consists of C, a sensitive (0.7 lx) charge-coupled device camera with lens removed, video amplifier VA, and video cassette recorder VCR. Synchronization of the optical pulses and camera read-out is accomplished with the aid of the end-of-field trigger (EOF), delay (DLY) and double pulser (DP). The divide by four circuit, (marked $\div 4$), prevents overlap of data from different pulse pairs.

diffuse a distance of $\lambda/4$, where λ is the wavelength of light. In a dense system where multiple scattering dominates, the observed much shorter multiple-scattering phase-breaking time τ_m , corresponds to very much smaller motions of the scatterers.

Our samples consisted of 0.46 μm diameter polystyrene spheres in a 10% (v/v) suspension in distilled water. The average phase-breaking time for single scattering from these spheres in dilute suspension is $\tau_1 = 970 \mu\text{s}$, as found both from experiment¹⁴ and standard diffusion theory. In high concentration, clustering and other effects inevitably slow particle diffusion, so that the observed reduction in phase-breaking time due to multiple scattering must be considered as a lower limit.

We photographed the time evolution of the emergent wave using short optical pulse pairs of variable separation obtained from an Ar-ion laser operating at 514.5 nm. Energy deposition in the sample due to absorption was estimated to be less than 4×10^{-7} J per pulse, so that heating effects were negligible. Using suitable electronics, backscattered speckle patterns were recorded by a video system in which the speckle patterns from each member of a pulse pair were individually stored on video cassette. A schematic diagram of the apparatus is shown in Fig. 1. Of special interest may be the fact that we were able to store speckle images separated in time by only 10 μs in sequential two-field frames separated by 33.3 ms. This was accomplished by synchronizing the laser pulses with the read-out cycle of the video camera in such a way that the first member of an optical-pulse pair arrived immediately before read-out of a given field, and was therefore stored in that field, whereas the second optical pulse arrived after read-out and was stored in subsequent fields.

In Fig. 2 we display pictures that dramatically illustrate the phase-breaking process. Two snapshots of the emergent wave which are initially separated by a time much shorter than that required for phase breaking appear nearly identical (Fig. 2a, b). As the time separation grows to be of order τ_m significant evolution of the wave becomes apparent (Fig. 2c, d), and after a time longer than that needed for phase breaking the wave evolves into a new, uncorrelated configuration (Fig. 2e, f). Note that this complete loss of correlation has already occurred in one-tenth of the single-particle diffusion time. From quantitative analysis of the cross-correlation function of numerous pairs of pictures such as those shown here, we estimate that multiple scattering produces a drastic 25-fold reduction in phase-breaking time, and that τ_m is approximately 40 μs .

We also obtained data similar to Fig. 2 for a sample in which a high concentration (0.02 mol l⁻¹) of rhodamine 620 dye was added to the suspension to produce a very short absorption length of approximately 15 μm at the laser wavelength. This

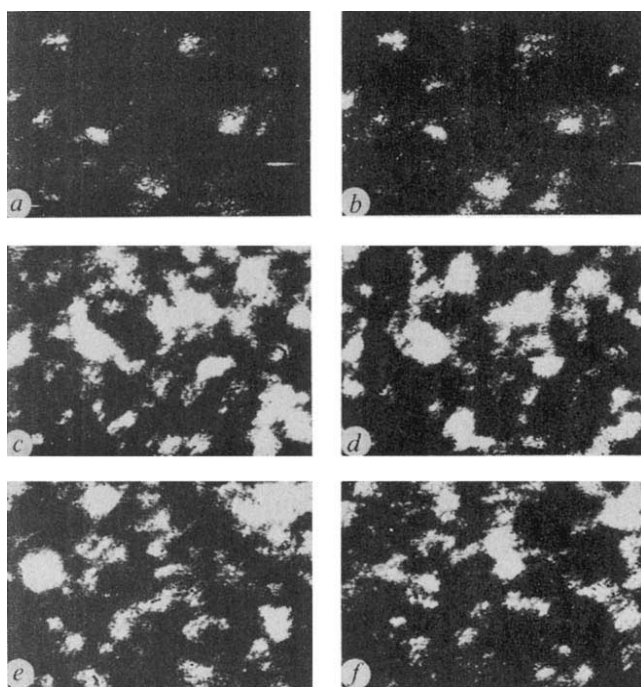


Fig. 2 Direct visual observation of phase breaking through snapshots of the time evolution of speckle patterns for a, b, a 10 μs optical pulse pair separated (centre-to-centre) by 15 μs ; c, d, a 20 μs pulse pair separated by 40 μs ; e, f, a 20 μs pulse pair separated by 100 μs . Digitization and calculation of the cross-correlation function, for numerous sets of picture pairs such as these, yield a multiple-scattering phase breaking time of $\sim 40 \mu\text{s}$.

sample showed a marked approximately fivefold increase in τ_m , demonstrating unequivocally that the observed reduction in phase-breaking time was not due to heating, because for this sample the energy deposition was 2×10^{-4} J per pulse, a factor 500 times larger than the energy deposited in the undyed samples.

Phase breaking occurs when scatterer diffusion changes the total optical path of length L of relevant photon trajectories by an amount approximately equal to $\lambda/4$. The projected mean distance δu along which a scatterer diffuses during a time τ is $\delta u = (2D\tau)^{1/2}$, where D is the diffusion constant of the scatterers. The total change ΔL in optical path for a photon trajectory consisting of N successive scattering events is $\Delta L \approx N^{1/2}\delta u$. Equating ΔL with $\lambda/4$ yields for the multiple-scattering phase-breaking time, $\tau_m \approx \tau_1/N$. The optical phase of long trajectories (large N) is thus seen to be a fragile quantity easily broken by small motions of the scatterers. Our results for the dyed sample are in full accord with these ideas, because restricting N by strong absorption produced the expected marked increase in τ_m .

The observed 25-fold decrease in phase-breaking time due to multiple scattering implies that photon trajectories with N much larger than 25 are irrelevant for backscattered universal optical fluctuations¹⁵⁻¹⁷. This new result would also appear to have significant implications for the weak localization of photons¹⁸⁻²¹, universal conductance fluctuations²², and low-dimensional^{23,24} and mesoscopic²⁵ systems. Further discussion of these matters will be given elsewhere.

We thank M. Hoshen for his assistance in performing the measurements and in carrying out the picture digitization and cross-correlation calculations.

Received 8 January; accepted 27 February 1987.

1. Anderson, P. W. *Phys. Rev.* **109**, 1492-1505 (1958).
2. Mott, N. F. *Metal Insulator Transitions* (Taylor & Francis, London, 1974).
3. Anderson, P. W. *Phil. Mag.* **B52**, 505-509 (1985).
4. John, S. *Phys. Rev. Lett.* **53**, 2169-2172 (1984).
5. Golubentsev, A. A. *JETP* **59**, 26-32 (1984).

6. Abrahams, E., Anderson, P. W., Licciardello, D. C. & Ramakrishnan, T. V. *Phys. Rev. Lett.* **42**, 673-676 (1979).
7. Gorkov, L. P., Larkin, A. F. & Khmel'nitsky, O. E. *JETP Lett.* **30**, 228-232 (1979).
8. Bergmann, G. *Phys. Rep.* **107**, 1-58 (1984).
9. Lee, P. A. & Ramakrishnan, T. V. *Rev. mod. Phys.* **57**, 287-337 (1985).
10. Mott, N. F. & Kaveh, M. *Adv. Phys.* **34**, 329-401 (1985).
11. Thouless, D. J. *Phys. Rev. Lett.* **39**, 1167-1169 (1977).
12. Alshuler, B. L., Aronov, A. G. & Khmel'nitsky, D. E. *Solid St. Commun.* **39**, 619-623 (1981).
13. Mott, N. F. & Kaveh, M. *Phil. Mag.* **B52**, 177-183 (1985).
14. Schumacher, R. T. *Am. J. Phys.* **54**, 137-141 (1986).
15. Etemad, S., Thompson, R. & Andrejco, M. J. *Phys. Rev. Lett.* **57**, 575-578 (1986).
16. Kaveh, M., Rosenbluh, M., Edrei, I. & Freund, I. *Phys. Rev. Lett.* **57**, 2049-2052 (1985).
17. Shapiro, B. *Phys. Rev. Lett.* **57**, 2168-2171 (1986).
18. Kuga, Y. & Ishimaru, A. *J. opt. Soc. Am.* **A1**, 831-835 (1984).
19. Tsang, L. & Ishimaru, A. *J. opt. Soc. Am.* **A2**, 2187-2193 (1985).
20. Van Alba, M. P. & Legendijk, A. *Phys. Rev. Lett.* **55**, 2692-2695 (1985).
21. Wolf, P. E. & Maret, G. *Phys. Rev. Lett.* **55**, 2696-2699 (1985).
22. Lee, P. A. & Stone, A. D. *Phys. Rev. Lett.* **55**, 1622-1625 (1985).
23. Pepper, M. *Contemp. Phys.* **26**, 257-293 (1985).
24. Kaveh, M. *Phil. Mag.* **B52**, 521-540 (1985).
25. Imry, Y. in *Directions in Condensed Matter Physics* (eds Grinstein, G. & Mazenko, E.) 275-293 (World Scientific, Singapore, 1986).

Vertical distribution and alteration of dikes in a profile through the Troodos ophiolite

J. M. Hall*, B. E. Fisher*, C. C. Walls*, S. L. Hall*,
H. P. Johnson†, A. R. Bakor‡, V. Agrawal§,
M. Persaud|| & R. M. Sumaiang¶

* Centre for Marine Geology, Dalhousie University, Halifax, Nova Scotia B3H 3J5, Canada

† Department of Oceanography, University of Washington, Seattle, Washington 98195, USA

‡ Faculty of Earth Sciences, King Abdul Aziz University, PO Box 1744, Jeddah, Saudi Arabia

§ Geophysics Division, Geological Survey of India 23-C Gokhale Marg, Lucknow 226001, India

|| Guyana Geology and Mines Commission, PO Box 1028, Georgetown, Guyana

¶ Department of Minerals and Energy, Geological Survey, PO Box 778, Port Moresby, Papua New Guinea

In situ oceanic crust at intermediate depths of from one to five km is presently largely inaccessible to continuous examination and sampling except in the possibly special circumstances of major transform faults. The study of ophiolites over this interval, where thin, sub-vertical dikes are usually important, may provide guidance to the nature of *in situ* oceanic crust at these depths. The likelihood that similarities exist between *in situ* and ophiolitic oceanic crust has recently been strengthened by the possible identification of a dike complex below 0.8-km depth in DSDP Hole 504B¹. Sub-vertical basic dikes occur over a 4.6-km depth interval from 0.2 to 4.8 km crustal depth in a profile through the Troodos, Cyprus, ophiolite. Exponential growth and arctangent decay in dike density occur over 2.3-km and 1.0-km intervals, respectively. Secondary magnetite is the dominant dike magnetic material of the dikes and associated screens. These results have implications for the construction, alteration history and magnetization of *in situ* oceanic crust at intermediate depths.

The profile is located in the principal Cyprus Crustal Study Project (CCSP) area (Fig. 1) and has been compiled using a combination of drill-core and outcrop information together with the assumption of a sinusoidal model for the regular periclinal emplacement structure of the ophiolite (Fig. 2). The justification for linking directly vertical, drill-core and horizontal, outcrop measurements of dike density is given in ref. 2.

The interval over which dikes occur can be divided into five zones. Occasional dikes occur in the first zone between 0.2 and 0.5 km. A sharp rise in dike abundance, from a few per cent to 20-40% of the section, occurs in a second zone, extending from about 0.5 to 0.7 km. The third zone encompasses exponential

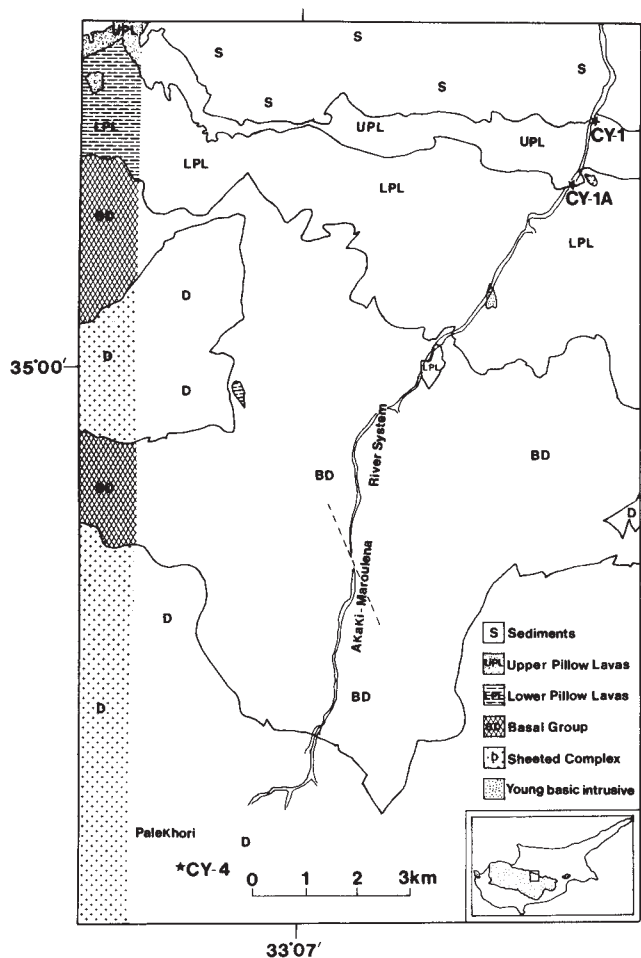


Fig. 1 Location map for Cyprus Crustal Study Project principal study area. *, Drill sites, with numbers. Note major crush zone (---) and roof of late (young) basic intrusive along line of Akaki-Maroulina River System; Insert: Cyprus, Troodos ophiolite stippled, map area indicated by rectangle.

growth in dike abundance from 20-40% to 100% of the section over the 2.3-km interval from 0.7 to 3 km according to:

$$F = 0.11 e^{0.68d}, r^2 = 0.77 \quad (1)$$

(where F is the fraction of dikes in a horizontal profile transverse to the strike of dikes, and d is the crustal depth in kilometres). Significant departure from this expression only occurs for $F > 0.8$, where sheeted dike conditions are approached more rapidly than is predicted by the expression. Both observations and (1) show the crustal depth at which dikes comprise half the section is close to 2.2 km, while increase in F from 0.25 to 0.75 occurs over the 1.75 km interval from 1.0 to 2.75 km. The balance of this third zone consists of screens of extrusive material. At relatively shallow depths, screens consist of coherent, easily recognizable, short segments of pillowed or sheet flows, or of brecciated material. At depth in the interval, screens often consist only of mineralized, brecciated material.

The fourth zone, in which $F = 1.0$, is partly represented in outcrop and partly in the upper part of Hole CY-4. The outcrop thickness of this zone is difficult to evaluate, since this area is within the crestal area of the Troodos pericline where the modelled dip is close to the topographic gradient. A figure of 0.5 km for the surface exposed area is assumed here but this may be uncertain by as much as several hundred metres. In contrast, the part of this zone in Hole CY-4 is quite definite, the first screen of plutonic material occurring at 250 m, or 3.75 km crustal

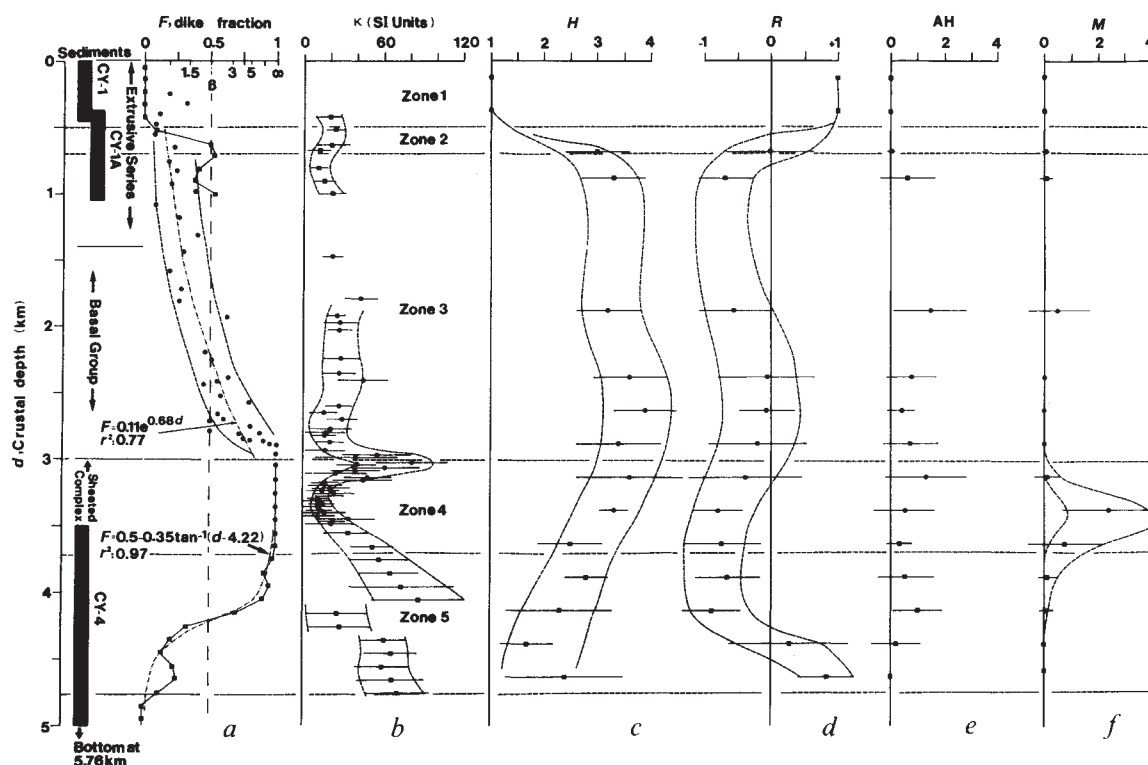


Fig. 2 Distribution of dike abundance, dike initial susceptibility and dike Fe-Ti oxide alteration properties with depth in the Cyprus Crustal Study Project area. *a*, Geological section: compiled from CCSP drillhole data and 1:31,680 Akaki-Lythrodonda map sheet included with⁸. Dike abundance profile. ■, Drillhole data; ●, outcrop data. *b*, Dike susceptibility profile (units are 10^{-3} SI). Values shown are, for drill-core intervals, averages for 100-m depth intervals and, for outcrop intervals, averages for each profile measured. Bar length indicates ± 1 standard deviation about averages. Drillhole interval average values are based on from 4 to 990 individual measurements, depending on the fraction of dikes in each 100 m interval. Profiles measured on outcrop average 60 individual measurements. *c*, Dike Fe-Ti oxide hydrothermal alteration index, *H*. Based on a 1 (no alteration) to 6 (complete dissolution of Fe from grains) degree of alteration scale. Points for this and other oxide properties are average ± 1 standard deviation of within sample average values for 250-m depth intervals. On average ten sample values contribute to each point. *d*, Relative abundance of primary and secondary magnetite, *R*, ratio is given by

$$\frac{P-S}{P+S}$$

where *P* and *S* are measures of the contents of altered primary and secondary magnetite respectively. *e*, Patchy haematitic alteration of secondary magnetite, *AH*, and *f*, martitic alteration of secondary magnetite, *M*. The scales are for abundance of alteration, 1, alteration absent; 4, alteration present to a significant degree in all grains.

depth. Thus, this zone is taken as extending from 3.0 to 3.75 km and has a thickness of 0.75 ± 0.25 km, the uncertainty relating to the outcrop part of the zone.

The fifth zone comprises the transition from the Sheeted Complex to the Plutonic Complex. This extends over 1.0 km from 3.8 to 4.8 km and the form of the decay in dike density can be closely modelled by a symmetrical arctangent expression:

$$F = 0.5 - 0.35 \tan^{-1}(d - 4.22), r^2 = 0.97 \quad (3.5 \leq d \leq 5.0) \quad (2)$$

Both observations and (2) show that the depth at which dikes comprise half the section is 4.2 km and the decay in *F* from 0.75 to 0.25 occurs over the 0.2-km interval from 4.1 to 4.3 km.

Alteration of the dikes is assessed here from a combination of oxide petrographic and magnetic properties.

Neither of the two dikes at shallow depth in the extrusives (zone 1) show any oxide-petrographic indication of alteration. Thermo-magnetic cycling indicates the development of moderate cation deficiency in titanomagnetite, indicative of a mild low temperature alteration that is frequently difficult to detect optically.

The sharp rise in dike density seen in the second zone is matched by the onset and sharp increase in hydrothermal alteration³ in both dikes and associated screens, which is present at intermediate levels throughout this zone. The diagnostic features

of this type of alteration are the occurrence of $1 \mu\text{m}$ anatase granules in magnetite, and the partial replacement by sphene of ilmenite lamellae formed during deuteric oxidation⁴ (Fig. 3a). The first indications of the development of secondary magnetite⁵ either as overgrowths on hydrothermally altered primary grains or in partly reconstructed primary grains (Fig. 3b) occurs in both the dikes and screens of this zone. This ratio of primary to secondary magnetite shows a sharp downward decrease from +1 (all primary magnetite) to -0.7 (largely secondary magnetite) that is closely coincident with the rapid downwards increase in dike density in this second zone. In both this transition and the deeper transition of zone 5 associated with reduction in dike density, secondary magnetite becomes dominant where $F \geq 0.3$. Haematitic (Fig. 3c) and martitic (Fig. 3d) alteration of secondary magnetite make their appearance in this zone. With the exceptions noted below, thermo-magnetic analysis for this and subsequent zones yields single 'magnetite' (525 to 575 °C) Curie points from largely reversible cycles. The 2.3 km transition to zone 4 conditions is characterized by maintenance of intermediate average degrees of hydrothermal alteration of primary titanomagnetite. Secondary magnetite is the dominant form of magnetite in this zone. Patchy haematitic alteration of secondary magnetite occurs irregularly throughout the zone. Martitic alteration is present but is less common than

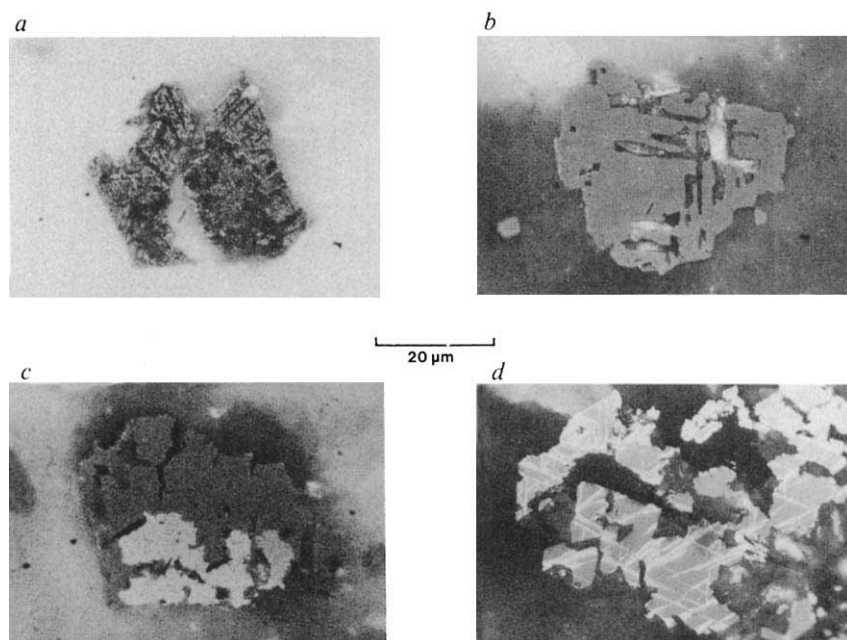


Fig. 3 Significant types of alteration of Fe-Ti oxides in the profile. *a*, Hydrothermal alteration of primary titanomagnetite (grey), $H = 4$. Ilmenite lamellae produced during deuteric oxidation partly replaced by sphene (black). General corrosion of grain is also the result of hydrothermal alteration. Minor reconstruction with localized formation of secondary magnetite, 0.71-km crustal depth. *b*, Largely reconstructed magnetite. Uniform secondary magnetite (light grey), elongate and rhomboidal sphene masses and holes represent the relics of a hydrothermally altered ilmenite lamellae system, 3.09 km crustal depth. *c*, Partial replacement of secondary magnetite (light grey) by haematite (off white), 3.01 km crustal depth. *d*, Martitic texture—haematite lamellae in secondary magnetite, 3.48 km crustal depth.

patchy alteration to haematite. Dike magnetic susceptibility, κ , which has shown little variation about an average value of $25\text{--}30 \times 10^{-3}$ SI units, rises sharply to about twice this value over the lowest 0.2 km of zone 3. The average state of hydrothermal alteration reaches a maximum at close to the zone 3–zone 4 boundary, below which it decreases to the bottom of zone 5. The relative abundance of secondary magnetite increases downwards in the upper part of the zone to reach average ratio values of ~ 0.7 again over the lower half of zone 4 and the upper part of zone 5. The replacement of secondary magnetite by haematite occurs at moderate levels throughout the zone while the occurrence of martitic alteration of secondary magnetite shows a sharp peak within the lower half of zone 4. Since both drill-core and outcrop samples contain martite and secondary haematite, these phases must have formed during original crustal construction and alteration rather than during recent sub-aerial weathering. Ubiquitous magnetite Curie points in this zone are consistent with the widespread presence of secondary magnetite, while susceptibility shows magnetite abundance to be at a minimum in the central part of the zone with regular increase towards both margins. Within zone 5 all alteration indicators show a downward decrease, with most of the reduction in secondary magnetite abundance taking place in the short interval centred at 4.2 ± 0.1 km depth where F falls from 0.75 to 0.25. Below this depth, primary Fe-Ti oxides are again commonly seen. The reduction in alteration with depth results in the presence below 4.5 km of some dikes in which the oxide and magnetic properties are the same as those commonly found in extrusives from the upper 0.5 km of *in situ* and Troodos oceanic crust. For these dikes, only low temperature alteration, resulting from contact with water at $< 50^\circ\text{C}$, has occurred. The increase with depth in oxide abundance in the lower part of zone 4, as indicated by increase in k , is continued strongly to 4.2 km depth in zone 5. Below this depth a 0.2-km low is followed by 0.5 km of values at close to the maximum reached at 4.2-km depth.

In terms of crustal construction basic dikes occur with strongly asymmetric distribution over 80% of the interval from the sediment–volcanic interface to 2 km below the first occurrence of plutonics, and comprise more than half the section over 2 m (35%) of this interval. The Troodos profile closely follows that

of DSDP Hole 504B to 0.75 km depth, below which the apparent difference between the profiles is difficult to confirm as a result of poor core recovery from the DSDP hole. The sequence of alteration stages in the Troodos profile suggest increased temperature and decreased fluid availability at depth and with time during crustal construction. The widespread occurrence of secondary magnetite implies that remanent magnetization is not of initial cooling origin over 3–4 km of the profile. The apparently uniform polarity of both primary and secondary magnetite bearing units could be either the result of crustal formation during the 25 Myr long normal polarity interval between 83 and 118 Myr^{6,7} or hydrothermal alteration and secondary magnetite formation rapidly following dike intrusion. Magnetite content, one control on remanence intensity, is uniform at 1.0–1.3% over the upper 2.5 km of the profile but shows a series of strong, continuous variations in the lower part of the profile, of which the trough within zone 4 may be the result of the substantial replacement of magnetite by haematite.

This study was funded by NSERC (Canada) through grants A4652, A9917, G1563, A7812, and IDRC (Canada) grant 3-P-81-1012. We thank Jane Barrett and Ference Stefani for help in the preparation of the manuscript. The drilling operations at Holes CY-1, CY-1a and CY-4 were supported financially by the Governments of Cyprus and Canada, grants from IDRC (3-P-81-1012), NSERC (G0741 and A4652), and Dalhousie University, NERC (UK) grant GR3/4473, SNF (Denmark) grant 11-11732, the Volkswagen Foundation (FRG) grant I/38 488, NSF (USA) grant EAR83-11776, and the Faculty of Earth Sciences, King Abdul Aziz University, Jeddah, Saudi Arabia.

Received 25 September 1986; accepted 3 February 1987.

1. Anderson, R. N. *et al.* *Nature* **300**, 589–594 (1982).
2. Hall, J. M. & Walls, C. C. *Prog. Abstrs. Geological Association of Canada/Mineralogical Association of Canada Ann. Mtg.* **11**, 76 (1986).
3. Ade-Hall, J. M., Palmer, H. C. & Hubbard, T. P. *Geophys. J. R. astr. Soc.* **24**, 137–174 (1971).
4. Haggerty, S. E. in *Oxide Minerals 3* (ed. Rumble, D.) Hg-1-Hg-100 (Miner. Soc. Am. Short Course Notes, Washington DC, 1976).
5. Hall, J. M. *Can. J. Earth Sci.* **22**, 85–101 (1985).
6. Harland, W. B. *A Geologic Time Scale* (Cambridge University Press, 1982).
7. Blome, C. D. & Irwin, W. P. *Geology* **13**, 401–404 (1985).
8. Bear, L. M. Geological Survey Department Cyprus Memoir No. 3 (Government of Cyprus, Nicosia, 1960).

Possible resonance effect in the distribution of earthquake damage in Mexico City

J. Flores, O. Novaro & T. H. Seligman

Instituto de Física, UNAM AP 20-364, México 01000 DF, México

The destruction caused by the earthquake of 19 September 1985 in Mexico City shows three remarkable features¹. First, a concentration of damage in the former lake bed; second, a peculiar distribution of high- and low-damage areas alternating within a few city blocks, and third, a selectivity for buildings between five and fifteen storeys high. While the last point is understood by civil engineers² and the first confirms the relevance of the soft ground, the second feature is the most puzzling. To a physicist such a pattern evokes the idea of a standing wave, with low- and high-damage areas corresponding to nodes and antinodes. This suggests a resonance phenomenon of the waterlogged ground moving collectively. From the Navier equations of classical elasticity, it is known³ that a transfer of the incoming seismic-wave energy to longitudinal waves takes place if a sharp change from hard to soft ground exists. Solving then the two-dimensional Poisson equation for the longitudinal waves within the ancient lake boundaries, we show that for the observed seismic wave period of 2 s, a standing-wave solution exists which explains qualitatively the damage distribution pattern.

It has long been established⁴ that within the valley of Anahuac, in the area presently occupied by Mexico City, there are three types of terrain: the firm ground of the foothills and lava beds, the so-called transition area, and the soft ground that corresponds to the old lake where the city area is located. This soft ground still has a very high water content (~90% in volume and 75% in weight^{4,5}) and the effects of this for building foundations is known to be perilous⁶. The soft-ground area boundary was first determined by Marsal and Mazari⁴ and further studied by Reséndiz *et al.*⁷. It is shown in Fig. 1 to which we shall refer all our calculations.

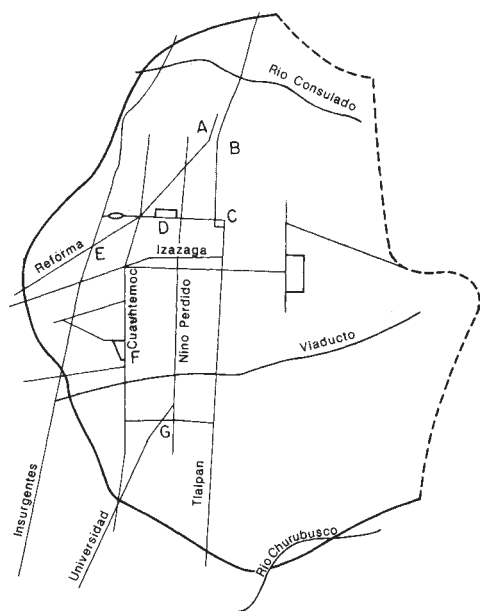


Fig. 1 Boundaries of the soft-ground region as determined by Marsal and Mazari⁴. The eastern borders were adjusted following a personal suggestion of Marcos Mazari. Some city streets and areas are marked for easier reference including: A. Tlatelolco, B. Tepito, C. Zócalo, D. Alameda Park, E. Intersection of Reforma and Insurgentes Boulevards, F. Centro Médico and G. SCT Building.

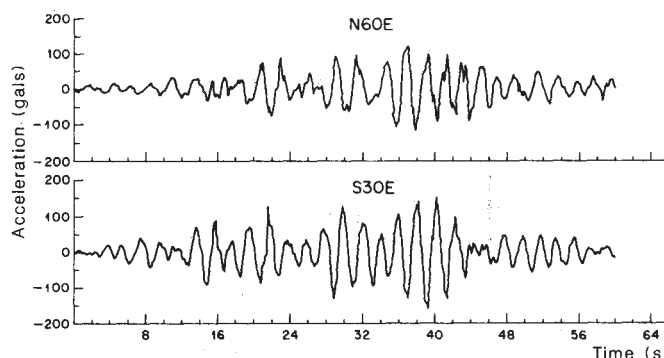


Fig. 2 Accelerogram obtained by Instituto de Ingeniería at the SCT building⁹ showing accelerations of 15 gals or more (the maximal value reported in ref. 2 is 0.2g), as well as a rather well-defined period of two seconds and a prolonged duration for the seismic movement of 19 September 1985. The accelerograph is triggered by two successive wave arrivals due to a difference of the P- and S-wave velocities¹⁷. The former are a factor 1.6 to 2.0 times faster¹⁸. This is evident from the accelerogram.

The effect that this soft ground can have in augmenting the dangers during the frequent earthquakes which originate from displacements of the Cocos and Central American Plates was dramatically brought to public attention during the events of 19 September 1985. Within this soft-ground boundary, or in its southeastern extension corresponding to old water-way canals linking the lakes of Tenochtitlan and Xochimilco, lies well over 95% of the buildings that collapsed during the prolonged (2 min) and violent (8.1 Richter scale) movement of a quite distant (almost 400 km away) section of the Cocos plate.

The acceleration of the ground was measured by the accelerographs of the Instituto de Ingeniería located in different areas of the city⁸⁻¹⁰. In the lava beds to the south of the valley an acceleration of 0.07g (*g* is the acceleration due to gravity) was registered⁸. The accelerograph⁹ located in the garden of the SCT building in the soft-ground region however, gave a substantially higher value of 0.20g, as can be seen in Fig. 2. Also very conspicuous in this figure is the remarkable monochromaticity of the oscillations, which show a dominant period of two seconds. This quite well-defined period explains the fact mentioned above that buildings of intermediate height, whose own natural vibrational periods are near 2 s, were by far the most severely hit in the city area.

One aspect of the recent earthquakes that is not well understood is the surface distribution of the damage within the soft-ground area depicted in Fig. 1. From existing reports¹¹ we can note that whole city blocks with heavily destroyed buildings alternate with other blocks of similar buildings that escaped unscathed. Some of the neighbourhoods that were severely affected last September had historically been prone to collapses in documented earthquakes over past centuries¹². Analogous patterns of damage, although with significant variations, can be established for recent seisms originating in different epicentres¹. This leads to the question of whether the response of the soft-ground region of Fig. 1 as a whole implies a pattern of maximal and minimal ground accelerations that could explain the main disaster zonal distribution.

The Mexico City tremors can be modelled roughly by assuming that a seismic wave travels in a rigid semi-infinite medium surrounding the lake bed. In this case it can be shown¹³ that the solutions of the Navier equations can be separated into transverse and longitudinal components, each one of which obeys a scalar Poisson equation, the coupling between them occurring only at the boundary separating the two different media. When one of the media is rather hard and the other quite soft, however, a considerable amount of the energy is transferred from transverse to longitudinal modes, as is well known^{3,14,15}. By using the adequate parameters for the hard- to soft-ground

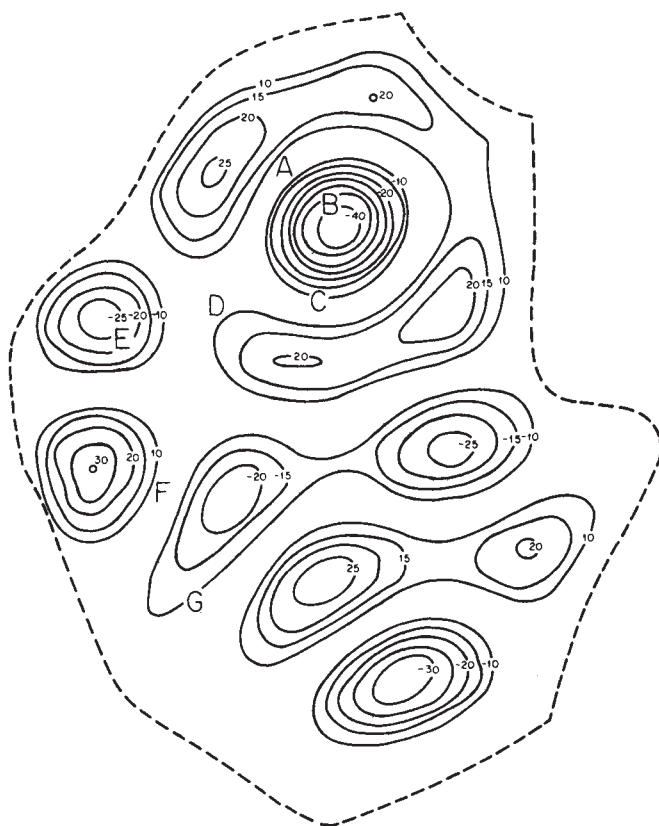


Fig. 3 Maxima and minima of the standing wave solutions of the Poisson equation showing the areas of the soft-ground region that present maximal displacement. Both maxima and minima correspond to largest accelerations according to our model. Many areas of extreme destruction in the Mexico City earthquake of 19 September 1985 are located in these maximal acceleration zones.

transition in the Valley of Mexico one confirms that S to P conversion is very important: $\sim 30\%$ of the incident S-wave energy is transmitted as longitudinal waves to the basin for angles of incidence of 45° or larger. As a matter of fact, we have obtained energy conversion ratios similar to those reported in Fig. 3.14 of ref. 3.

The transmitted transverse component is not negligible either, but this is of no consequence in the establishment of the destruction pattern we are dealing with. This is due to the small velocity of these S waves^{16,17}, which in the soft ground is $\leq 3\%$ ¹⁸ of the longitudinal wave velocity, which in turn is essentially the same as the speed of sound in water. S waves therefore do not have enough time during the 2-min earthquake to suffer the multiple reflections in the lake boundaries to produce the interference pattern. The transverse components then contribute uniformly over all of the lake region, adding to the maximal damage zones (the antinodes) and causing the nodes to be zones of lower damage rather than damage-free zones.

On the other hand, longitudinal waves in the soft ground have the speed required to establish an interference pattern and there is enough energy in the antinodes so that the actual zone damage distribution observed after the earthquake of 19 September 1985 may be explained.

We therefore pose the following theoretical problem: we solve the two-dimensional Poisson equation governing longitudinal waves in an elastic medium surrounded by perfectly rigid boundaries. The input data consist of the velocity of these waves, which is taken to be that of the soft ground of Mexico City (1.5 km s^{-1}); the boundaries are those of the ancient lake bed, whose shores are $\sim 9 \text{ km}$ long; and the frequency is that

measured during the September 1985 earthquake. With these data, the standing-wave solutions of the Poisson equation were obtained by discretizing the problem on a 250-m mesh and using the power method and a computer program provided by B. Neuberger (B. Neuberger and D. W. Noid, personal communication and ref. 19). The eigenvalues obtained from the program are the frequencies of vibration; the eigenvectors correspond to the ground displacements (and consequently the ground accelerations). We choose among the many solutions first by taking only those that have the correct frequency; then, among the few eigenvectors whose eigenvalues are clustered around the 2-s period, we select the one that most closely maintains the original direction of the incoming travelling wave to the Valley of Anahuac from the epicentre, which lies in a 30° south-east direction from Mexico City. In Fig. 3 we give an eigenvector which fulfils all these requirements. It shows a series of nodes and antinodes that originate from destructive and constructive interference of waves as they are reflected in the rigid boundaries of Fig. 1. The separation between the areas of maximum displacement reflects the wavelength of the standing waves, which is $\sim 3 \text{ km}$, implying an alternation of maximal with minimal damage zones separated by a characteristic distance of 750 m, in agreement with observed facts. Interestingly enough, the superposition of the eigenvector of Fig. 3 on to the damaged zones in Mexico City shows some very striking coincidences (notably with the heavily destroyed¹¹ Roma, Reforma-Insurgentes, Tepito, as well as other areas), as shall be discussed elsewhere.

We should note here that the rigid boundary conditions used to derive Fig. 3 do not correspond to the actual acoustic boundary conditions for a near liquid that would most faithfully represent the real medium of the lake bed. The relative positions of the antinodes (maxima and minima) should however be essentially equivalent for both boundary conditions because they are completely imposed by the geometry. This can be trivially demonstrated to be true for particularly simple geometries (for example, circular or rectangular) for which these maxima and minima (sinusoidal functions for the rectangular, Bessel for circular) for both rigid and acoustic boundary conditions are at coincident positions. The nodal structure is different for each condition of course, because for the acoustic case we have nodal points while for the rigid case we have lines. As concerns our main interest here, that is, the correlation of the standing wave solutions of the elastic medium with the most heavily damaged areas however, the nodes have no significance because as we have seen they are related to only one (the stationary longitudinal component) among several components of the seismic waves. We are presently calculating the acoustic conditions for the precise lake boundary of Fig. 1 in order to predict the nature and directions of movement of the ground for each specific zone within the city. These latter aspects are beyond the simple predictions that Fig. 3 can give and hopefully we may make further tests and comparisons between our model and the observed data of the September 1985 tremors.

We thank M. Mazari, F. Sánchez-Sesma, S. K. Singh and S. de la Cruz for their interest and valuable comments and through them the Instituto de Ingeniería and Geofísica of UNAM for important data and information. We also thank many of our colleagues at Instituto de Física for their criticism and encouragement, especially R. Pérez-Pascual, T. Brody and B. Lesche. Also, we thank F. del Rio and M. Chávez of Universidad Autónoma Metropolitana, Iztapalapa for their experiments on the velocity of L waves in clay.

Received 25 August 1986; accepted 22 January 1987.

1. *Información Científica y Tecnológica* Vol. 7, no. 110 (Consejo Nacional de Ciencia y Tecnología, México, 1985).
2. Anderson, I. *New Scientist* **108**, 32–33 (1985).
3. Ewing, W. M., Jardetsky, W. S. & Press, F. *Elastic Waves in Layered Media* 74–124 (McGraw-Hill, New York, 1957).
4. Marsal, R. J. & Mazari, M. *The Subsail of Mexico City* (UNAM, México, 1959).

5. Marsal, R. J. in *International Clay Conference*, Report E16 (Instituto de Ingeniería, UNAM, México, 1975).
6. Hiriart, F., Marsal, R. J. & Rosenbluth, E. *Ingeniería* **28**, 1-34 (1957).
7. Reséndiz, D. et al. *Inst. Ingen. UNAM*, Report No. 276 (1970).
8. Prince, J. et al. *Inst. Ingen. UNAM*, Report No. IPS-10A (September 20, 1985).
9. Mena, E. et al. *Inst. Ingen. UNAM*, Report No. IPS-10B (September 21, 1985).
10. Quaa, R. et al. *Inst. Ingen. UNAM*, Report No. IPS-10C (September 23, 1985).
11. Jaime, A. *Archivo de Edificios Dañados por los Sismos del 19 y 20 de Septiembre 1985 en la Ciudad de México* (Sección de Geotermia II UNAM, México, 1985).
12. Manzanilla, L. *Report Inst. Invest. Antrop. UNAM* **1**, 1-8 (1986).
13. Landau, L. D. & Lifshitz, E. *Theory of Elasticity* (Pergamon, London, 1959).
14. Aki, K. & Richards, P. G. *Quantitative Seismology* (Freeman, Berkeley, California, 1980).
15. Kolsky, H. *Stress Waves in Solids* (Dover, New York, 1983).
16. Love, A. *A Treatise on the Mathematical Theory of Elasticity* (Cambridge University Press, 1892).
17. Leet, L. *Earth Waves* (Harvard Univ. Press, Cambridge, 1950).
18. Martínez, B., León, J. L., Rascón, O. I. & Villareal, A. *Inst. Ingen. UNAM*, Report No. 338 (1974).
19. Neuberger, B. & Noid, D. W. *Report of the Oak Ridge National Laboratory* (ORNL, Tennessee, 1983).

Expulsion of fluids from depth along a subduction-zone decollement horizon

ODP Leg 110 Scientific Party*

Results from Deep Sea Drilling Project Leg 78A at the Barbados Ridge suggested that high fluid pressures are found along the decollement zone separating the underthrusting Atlantic Ocean crust and the overthrusting Caribbean plate, but drilling failed to penetrate this major fault surface¹. On Leg 110 of the Ocean Drilling Program we cored through this fault surface, achieving the first penetration of a decollement separating two plates in a subduction zone. Here we report pore-water chemistry, temperature anomalies and structural observations which indicate that the deforming sediments de-water through fracture permeability associated with faults, and that the accretionary wedge and underthrust sediments comprise distinct fluid domains separated by a permeability barrier above the decollement zone. Methane-bearing fluids derived from a probable thermogenic source travel >25 km beneath the shallowly inclined decollement.

Subduction zones consume water-rich sediment which evolves to highly deformed rock of negligible porosity, requiring large-scale fluid expulsion². The consequent fluid flow and attendant high fluid pressures influence many aspects of subduction-zone geology, including the shape of the accretionary wedge³, thrust faulting and fold vergence^{4,5}, heat transport⁶, diagenesis and metamorphism⁷, and benthic biology⁸. Ocean Drilling Program (ODP) Leg 110 investigated the structural and hydrological effects of this transition from undeformed deep-sea sediment to disrupted strata in the subduction zone separating the Caribbean plate and the Atlantic oceanic crust.

The prime objective of Leg 110, penetrating the decollement, was realized at our initial location, Site 671 (Figs 1, 2). There we cored 500 m of Pleistocene to Lower Miocene imbricately thrust, offscraped calcareous mud and mudstone, a 40-m-thick decollement zone of Lower Miocene to Upper Oligocene scaly (sheared and flattened) mudstone, and 150 m of little-deformed underthrust Oligocene mudstone, marlstone and siltstone. Penetration of an unstable Eocene sand layer halted drilling before oceanic crust was reached.

A reference hole at Site 672, 6 km east of the deformation front, defined the characteristics of all units penetrated in the accretionary wedge in a relatively undisturbed state. A distinctive lithology and excellent fossil control at the top of the decollement zone at Site 671 allowed correlation with equivalent strata at the reference site. Anomalously high porosity and a probably correlative decrease in sediment strength characterize the 'future decollement' and could explain why failure occurred at this level. Low-angle shear zones, and normal and reverse faults suggest incipient failure here, 6 km seaward of the deformation front. A methane anomaly near the level of the 'future decollement' suggests, furthermore, that fluid may be advecting laterally at this level. The Middle to Upper Eocene sandy interval which stopped drilling at Site 671 was completely penetrated at the reference site. Low chloride and relatively high methane contents in pore water from sandy layers also suggest lateral fluid flow from beneath the accretionary wedge, similar to that inferred farther south along the Barbados Ridge⁹.

In addition to documenting initial accretionary history, Leg 110 drilled two sites, 12 and 17 km up-slope and west of the deformation front, to measure the continuing evolution of the offscraped sediments. Here, at Sites 673 and 674, sedimentary rocks are far more deformed than in sites near the deformation front, and include structural features similar to those occurring in sub-aerially exposed accretionary complexes^{10,11}. More remarkably, this intensity of deformation was achieved with only modest de-watering, resulting in sedimentary deposits with ~50% porosity, as compared to an initial average porosity of 54% for all sediments at the reference site (Site 672).

Any model of the hydrogeology of an accretionary wedge must account for the sources of fluids, their mode of transport, and the geometry of their conduits. The modest de-watering of the offscraped sequence can be accounted for by a simple response of the sediment to vertical as opposed to lateral tectonic loading. The underthrust portion of the incoming sediment, although undeformed, is progressively loaded by the westward-thickening accretionary wedge which would undoubtedly result in consolidation, overpressure and water loss. Evidence for thinning and probable vertical consolidation of the underthrust sediment sequence is apparent on a seismic line (Discovery 17, line 109) north of the Leg 110 transect¹². Similar thinning might be apparent along the drilling transect; however, basement irregularities and normal faulting preclude its recognition.

Sedimentary deposits in the Leg 110 area are dominantly fine-grained and of low permeability¹³, with the exception of the locally sandy interval in the underthrust sequence. The geochemical evidence presented below argues for de-watering through the decollement zone. However, a simple calculation using Darcy's law shows that fluid movement through the intergranular permeability of the ubiquitous fine-grained sediments would be ~1,000 times slower than the rate of underthrusting. Hence, water cannot escape up the decollement zone except along fracture permeability. Rhodochrosite-filled veins in the scaly fabric of the decollement zone suggest that fluid pressures were sufficiently high to allow these surfaces to open for fluid flow. Apparently flow through the fine-grained sediments of the accretionary wedge and decollement zone is dominated by fracture permeability, especially along faults. In contrast, the Middle

* J. Casey Moore (Department of Earth Sciences, University of California at Santa Cruz, Santa Cruz, California 95064, USA). Alain Mascle (Institut Français du Pétrole, 1-4 Avenue Bois-Preau, BP 311, 92506 Rueil Malmaison Cedex, France). Elliott Taylor (Ocean Drilling Program, Texas A & M University, 1000 Discovery Drive, College Station, Texas 77840, USA). Patrick Andreieff (BRGM, BP 6009, 45060 Orleans Cedex-2, France). Francis Alvarez (Borehole Research Group, Lamont-Doherty Geological Observatory, Columbia University, Palisades, New York 10964, USA). Ross Barnes (Rosario Geoscience Associates, 104 Harbor Lane, Anacortes, Washington 98221, USA). Christian Beck (Département des Sciences de la Terre, Université de Lille, 59655 Villeneuve d'Ascq Cedex, France). Jan Behrmann (Institut fuer Geowissenschaften und Lithosphärenforschung, Universitaet Giessen, Senckenbergstrasse 3, D6300 Giessen, FRG). Gerard Blanc (Laboratoire de Géochimie, Université Pierre et Marie Curie, 4 Place Jussieu, 75252 Paris Cedex 05, France). Kevin Brown (Department of Geological Sciences, Durham University, South Road, Durham DH1 3LE, UK). Murlene Clark (Department of Geology, LSCB 341, University of South Alabama, Mobile, Alabama 36688, USA). James Dolan (Earth Sciences Board, University of California at Santa Cruz, Santa Cruz, California 95064, USA). Andrew Fisher (Division of Marine Geology and Geophysics, University of Miami, 4600 Rickenbacker Causeway, Miami, Florida 33149, USA). Joris Gieskes (Ocean Research Division A-015, Scripps Institution of Oceanography, La Jolla, California 92093, USA). Mark Hounslow (Department of Geology, Beaumont Building, Sheffield University, Brook Hill, Sheffield S3 7HF, UK). Patrick McLellan (Petro-Canada Resources, PO Box 2844, Calgary, Alberta, Canada). Kate Moran (Atlantic Geoscience Centre, Bedford Institute of Oceanography, BP 1006, Dartmouth, Nova Scotia B2Y 4A2, Canada). Yuiro Ogawa (Department of Geology, Faculty of Science, Kyushu University 33 Hakozaiki, Fukuoka 812, Japan). Toyosaburo Sakai (Department of Geology, Faculty of General Education, Usunomiya University 350 Mine-machi Usunomiya 321, Japan). Jane Schoonmaker (Hawaii Institute of Geophysics, 2525 Correa Road, Honolulu, Hawaii 96822, USA). Peter J. Vrolijk (Earth Sciences Board, University of California at Santa Cruz, Santa Cruz, California 95064, USA). Roy Wilkens (Earth Resources Laboratory, E34-404 Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA). Colin Williams (Borehole Research Group, Lamont-Doherty Geological Observatory, Columbia University, Palisades, New York 10964, USA).

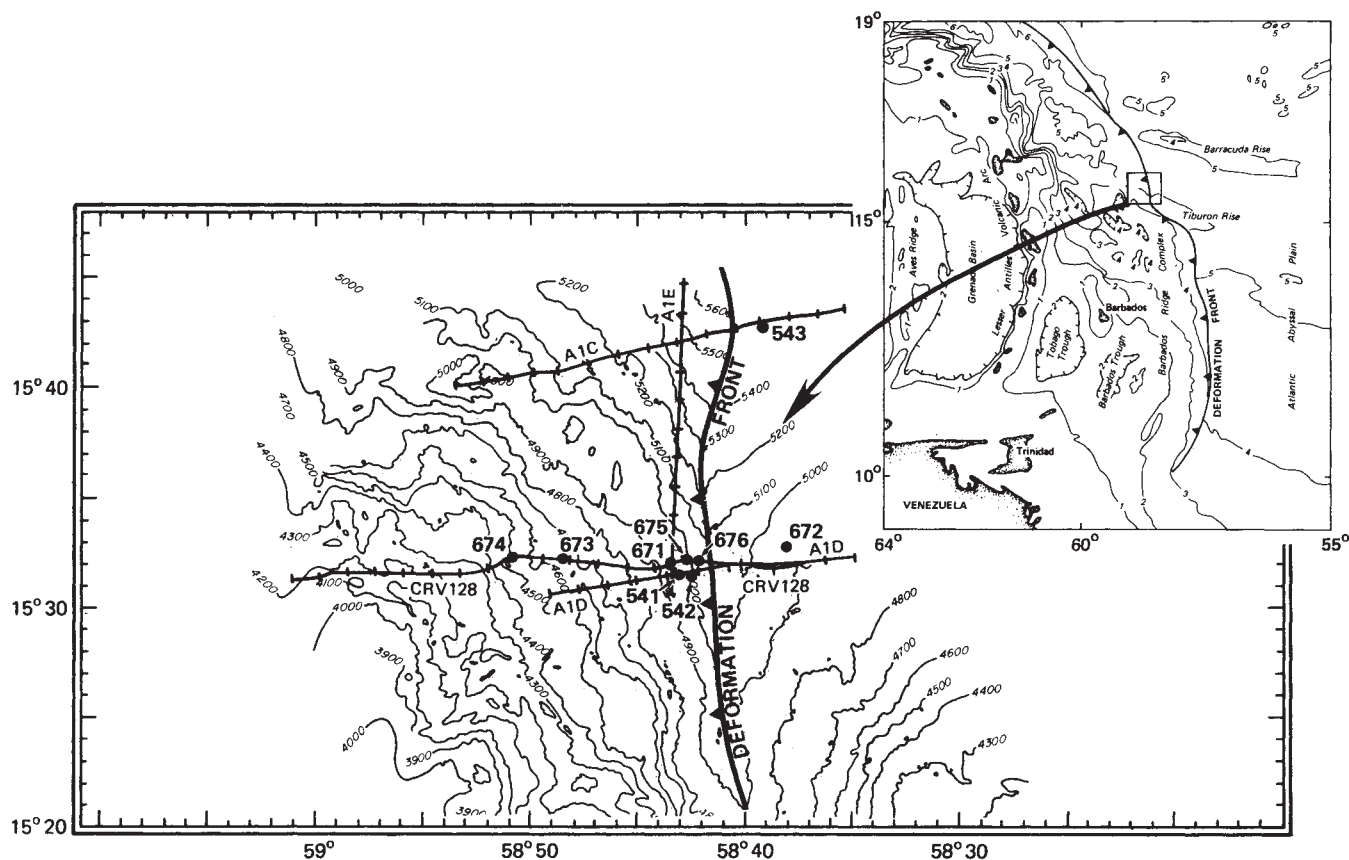
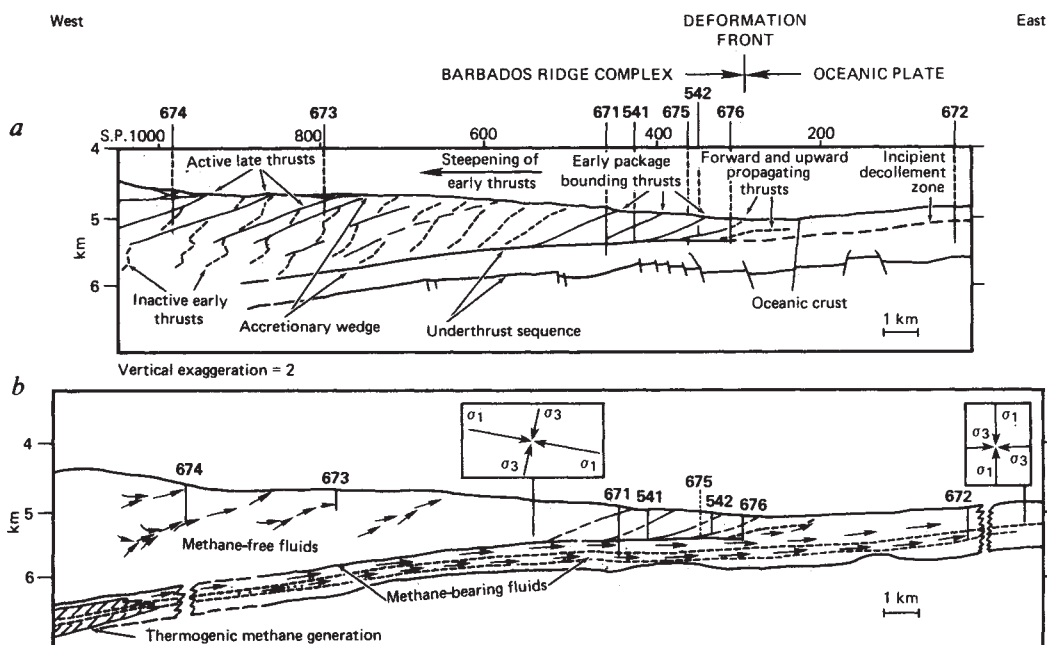


Fig. 1 Location map showing the setting of the eastern boundary of the Caribbean plate and the Atlantic abyssal plain (inset) as well as a detailed SeaBeam map with locations of the sites drilled during ODP Leg 110 and DSDP Leg 78A, and the network of seismic reflection lines.

Fig. 2 Cross-sections showing structural and hydrological framework of Leg 110 area. *a*, Structural framework of Leg 110 area based on a depth-section constructed from seismic line CRV 128 and all drilling results. At the deformation front offscraped sediment is imbricately stacked along young thrust faults; up-slope these faults and the deformed internal stratigraphy of the initial imbricate packages are cut by late thrust faults with an orientation similar to the early imbricate thrusts. *b*, Hydrological framework of Leg 110 area. Arrows show directions of fluid flow. Insets show state of stress in accretionary wedge and in oceanic sediment column well removed from disturbing effects of deformation front²¹. Because of decoupling along the decollement zone, the stress orientation in underthrust undeformed sediments should differ significantly from that in the overthrust accretionary wedge, perhaps more similar to their initial state in the abyssal plain.



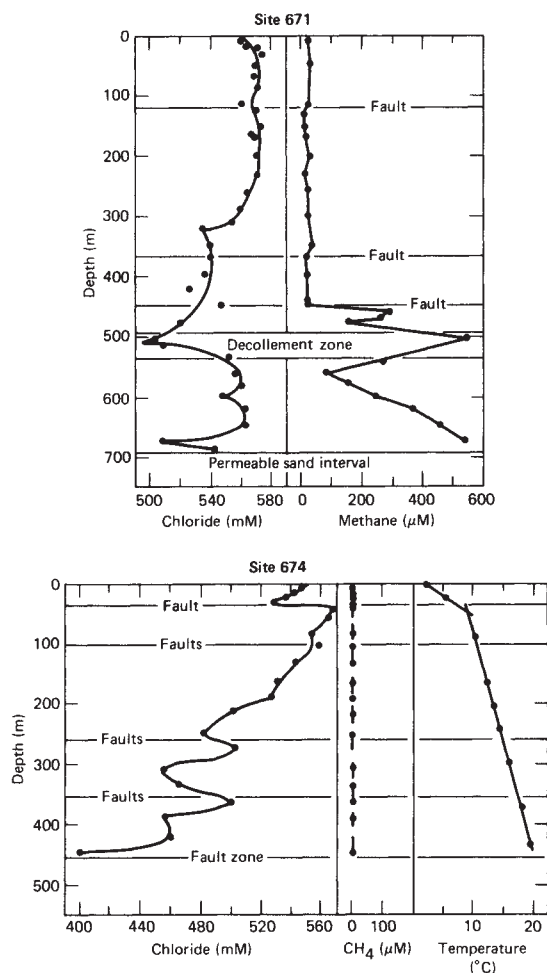


Fig. 3 Relationship of structural features, pore-water chemistry and temperature gradient. Low chloride and high methane at Site 671 correlate with decollement zone and top of permeable sand horizon and characterize the methane-bearing fluid regime. At Site 674 many, but not all faults correlate with lows in the chloride curve, suggesting some faults may be active conduits of fluid transport. The overall low-chloride values probably reflect a diffusive decrease of this constituent as a response to long-term flow of fluids through this highly faulted section. Low chloride and change in temperature gradient at 30 to 40 m sub-bottom at Site 674 suggests currently active flow. Note that methane is absent along all faults in the accretionary wedge at Site 674, characterizing the methane-free fluid domain. A small methane anomaly just above the decollement at Site 671 suggests minor leakage through this permeability barrier.

and Upper Eocene sandy deposits of the underthrust section probably transport fluid by intergranular flow through the permeable sands.

Chemical composition of the pore water provides the most clear-cut evidence for fluid movement along faults and, to a lesser degree, stratigraphic conduits. Many faults penetrated during Leg 110 included waters with low chloride contents (Fig. 3). Apparently, these low chloride values are caused by movement of fluid through a clay membrane which preferentially retains salts on the high-pressure side^{14,15}. In this ultra-filtration process the clays function like an osmotic barrier in retarding movement of salts. Accordingly, the faults are not only surfaces of fluid movement but must have, at times, a lower fluid potential than the surrounding sediments which are their fluid sources. The low chloride contents noted just above the sandy interval at Site 671 (Fig. 3) and within the same sequence at Site 672 suggest lateral flow along this permeable horizon. Because the observed pore-water anomalies would diffusively decay in $(1-2) \times 10^5$ yr¹⁶, they must represent active fluid flow. Anomalous

high concentrations of methane occur in pore waters within the decollement zone, in adjacent faults, and in the sandy interval below the decollement. But no methane anomalies are recognized in faults in the accretionary wedge more than a few tens of metres above the top of the decollement zone, even though these surfaces display low chloride values, indicating flow (Fig. 3). The accretionary wedge and underthrust sediment appear to constitute discrete fluid domains separated by the top of the decollement zone.

It is possible that the low-chloride waters observed in the Leg 110 cores could be due to the decomposition of methane hydrate, as occurs in the slope sediments of the Middle America Trench^{17,18}, rather than being due to ultra-filtration. But no hydrate reflectors appear in any of the seismic lines in the Leg 110 area; moreover, low chloride values occur both with and without associated methane. Hydrate decomposition produces both methane and low-chloride water.

Temperature data provide additional evidence supporting the geochemical arguments for fluid flow. High temperature gradients ($76-180^\circ\text{C km}^{-1}$) were measured at shallow levels at each site, implying unreasonably high temperatures at depth. At Site 674, where measurements were extended to several hundred metres sub-bottom, a near-surface gradient of $143^\circ\text{C km}^{-1}$ decreased to a more normal value of 28°C km^{-1} (Fig. 3), suggesting that the high surface gradients are explained by lateral heat input along faults. In fact, at Site 674 a change in temperature gradient, a negative anomaly in chloride, and a probable fault nearly coincide, indicating lateral flow of warm, low-chloride fluids along fracture permeability. A similar correlation of an increased temperature gradient up-section and a geochemical anomaly also occurred at Site 676, near the frontal thrust.

Methane at and below the decollement zone is probably not biogenically produced *in situ* because of the high local sulphate concentrations¹⁹. Instead, this hydrocarbon is probably advected through the underthrust sediment from beneath the accretionary wedge. As methane is not present in the upper levels of the accretionary wedge, it is not being produced biogenically here, notwithstanding low sulphate concentrations. Because the background geothermal gradient is $<30^\circ\text{C km}^{-1}$ at Site 674 and the accretionary wedge there is less than 2 km thick, methane is probably not generated thermogenically in adjacent accreted materials. Sediment below the decollement is deeply underthrust and would ultimately reach depths of thermogenic methane generation²⁰ west of Site 674, suggesting that the methane observed at Sites 671, 672, 675 has advected >25 km laterally.

From a simple geometrical perspective, fluids migrating from below the decollement should flow vertically to the sea floor rather than laterally along the shallowly inclined decollement. Material just over the decollement zone must present a permeability barrier to vertical flow. The preference for lateral flow may reflect the higher fracture permeability of the prominent scaly fabric along the decollement zone. Any permeability barrier over this zone may be caused in part by the expected change in stress orientation across the top of the decollement zone (Fig. 2b). The low plunge of the maximum principal stress in the accretionary wedge (Fig. 2a) would discourage the opening of steeply dipping fractures by high fluid pressure, while encouraging flow along shallowly dipping surfaces.

The ODP Leg 110 Scientific Party is grateful for the help extended to them by the ODP Technical and Engineering Groups, and the SEDCO personnel. The Ocean Drilling Program is sponsored by the NSF and participating countries under the management of Joint Oceanographic Institutions, Inc.

Received 7 October 1986; accepted 11 February 1987.

1. Biju-Duval, B. & Moore, J. C. *et al. Init. Rep. DSDP 78A*, 621 (US Government Printing Office, Washington, 1984).
2. Bray, C. J. & Karig, D. E. *J. geophys. Res.* **90**, 768-778 (1985).
3. Davis, D., Suppe, J. & Dahlen, F. A. *J. geophys. Res.* **88**, 1153-1172 (1983).
4. Hubbert, M. K. & Rubey, W. W. *Geol. Soc. Am. Bull.* **70**, 115-166 (1959).

5. Seely, D. R. in *Island Arcs, Deep Sea Trenches, and Back-Arc Basins* (eds Talwani, M. & Pitman, S. C.) 187-198 (Am. Geophys. Un. Maurice Ewing Series 1, Washington, 1977).
6. Langseth, M. G. & Hobart, M. A. *Eos* **65**, 1089 (1984).
7. Etheridge, M. A., Wall, V. J. & Vernon, R. H. *J. metamorphic Geol.* **1**, 205-226 (1983).
8. Kulm, L. D. *et al. Science* **231**, 561-566 (1986).
9. Westbrook, G. K. & Smith, M. J. *Geol.* **11**, 279-283 (1983).
10. Speed, R. C. & Larue, D. K. *J. geophys. Res.* **87**, 3633-3643 (1982).
11. Mascle, A., Biju-Duval, B., de Clarens, P. & Munsch, H. in *The Origin of Arcs* (ed. Wezel, D.) 375-400 (Elsevier, Amsterdam, 1986).
12. Westbrook, G., Mascle, A. & Biju-Duval, B. *Init. Rep. DSDP 78A* (ed. Biju-Duval, B.) 23-38 (US Govt Printing Office, Washington DC, 1984).
13. Marlow, M. S., Lee, H. & Wright, A. *Init. Rep. DSDP 78A* 549-558 (U.S. Govt Printing Office, Washington DC, 1984).
14. Freeze, R. A. & Cherry, J. A. *Groundwater* 604 (Prentice-Hall, Englewood Cliffs, New Jersey, 1979).
15. Kharaka, Y. K., Hull, R. W. & Carothers, W. W. in *Relationship of Organic Matter and Mineral Diagenesis* 79-174 (Soc. econ. Paleont. Miner., Short Course no. 17, Tulsa, 1985).
16. Gieskes, J. M. A. *Rev. Earth planet. Sci.* **3**, 433-453 (1975).
17. Hesse, R., Level, J. & Gieskes, J. M. *Init. Rep. DSDP 84* (eds von Huene, R. *et al.*) 727-736 (US Govt Printing Office, Washington DC, 1984).
18. Gieskes, J. M., Johnston, K. & Boehm, M. *Init. Rep. DSDP 84* (eds von Huene, F. *et al.*) 961-967 (US Govt Printing Office, Washington DC, 1984).
19. Claypool, G. E. & Kaplan, I. R. in *Natural Gases in Marine Sediments* Vol. 3 (ed. Kaplan, I. R.) 99-139 (Plenum, New York, 1974).
20. Waples, D. W. *Geochemistry in Petroleum Exploration 232/Inter.* (Human Resources Development Corp., Boston, 1985).
21. Davis, D. M. *Init. Rep. DSDP 78A* (eds Biju-Duval, B. *et al.*) 569-579 (US Govt Printing Office, Washington, 1984).

Species richness for parasitoids of British phytophagous insects

Bradford A. Hawkins & John H. Lawton

Department of Biology, University of York, York YO1 5DD, UK

Variation in the number of phytophagous insect species associated with different plant hosts has been examined for many insect and plant groups¹⁻⁷. Here we extend such analyses to focus on the number of parasitoids and hyperparasitoids attacking British phytophagous insects. First, we examine a number of variables describing characteristics of the herbivores and their host plants. Then we construct a multiple regression model to order the variables and determine how much of the variance in parasitoid species richness we can account for. We find that host insects' feeding niche, geographical extent of study ('range'), the architecture of plants on which hosts occur and the taxonomic isolation of host insects each significantly affects the number of associated parasitoid species per host species, but only the first two variables were significant in the multiple regression. The final model accounts for 22% of the variance in parasitoid richness.

Food chains comprising green plants, insect herbivores and parasitoids include over half of all known species of metazoans^{7,8}; hence, establishing the determinants of parasitoid species richness is a major step in understanding the diversity of terrestrial communities. We address this problem here using the parasitoids of British phytophagous insects. Parasitoid lists were compiled from primary literature sources, where authors reared species themselves or cited other primary records. The data include both the third and fourth trophic levels because many parasitoids are facultatively hyperparasitic and, as most hyperparasitoids have to locate the appropriate herbivore to oviposit, both are subject to similar ecological and evolutionary constraints. Table 1 summarizes the 285 species of herbivores from 42 families included in our analysis. The data and all sources will be published elsewhere.

We examined seven factors that we hypothesized should influence the number of parasitoid species attacking a species of phytophagous insect:

(1) Extent studied/range. We distinguished three levels in this variable: species whose parasitoids have been studied at a single site or in several sites in close proximity; in a region (for example, southern England or Wales); or throughout Great Britain. This variable represents a combination of 'entomologist-hours' and host insect range, but, as has been discussed elsewhere^{7,9}, such

Table 1 Taxonomic distribution of host insects included in the analysis

Order	Family	No. of genera	No. of species
Homoptera	Aphididae	10	10
	Cercopidae	2	2
	Cicadellidae	23	43
	Delphacidae	9	9
	Lachnidae	1	1
Hemiptera	Pemphigidae	1	1
	Miridae	7	9
Coleoptera	Apionidae	1	9
	Bruchidae	1	1
Diptera	Chrysomelidae	3	5
	Curculionidae	4	8
	Nitidulidae	1	1
	Scarabaeidae	1	1
	Scolytidae	4	6
	Agromyzidae	4	10
	Anthomyiidae	3	3
	Cecidomyiidae	14	29
	Chloropidae	1	2
	Opomyzidae	1	1
	Psilidae	1	1
	Tephritidae	2	4
	Tipulidae	1	2
	Arctiidae	1	1
Lepidoptera	Choreutidae	2	2
	Coleophoridae	1	5
	Eriocranidae	1	3
	Geometridae	7	13
	Gracillariidae	1	20
	Lycaenidae	2	2
	Lyonetiidae	1	3
	Nepticulidae	2	7
	Noctuidae	3	4
	Nymphalidae	2	2
	Oecophoridae	2	2
	Papilionidae	1	1
	Pieridae	3	5
	Satyridae	1	1
Hymenoptera	Tortricidae	13	14
	Yponomeutidae	2	2
	Cynipidae	11	25
	Eurytomidae	1	1
	Tenthredinidae	6	14
Total		158	285

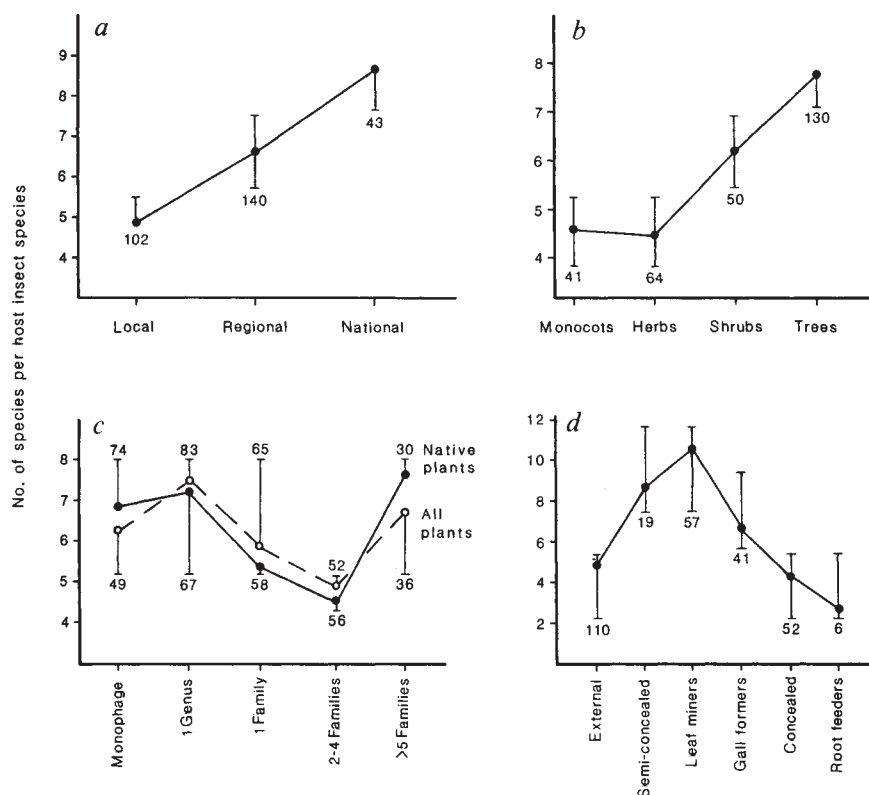
effects are implicit in all range data. We expected more extensively studied and/or more widespread insects to have more species of parasitoids.

(2) Architecture of herbivores' host plants. Food plants (L. K. Ward, personal communication) were classified into four groups: monocotyledons (including grasses), perennial dicotyledonous herbs, shrubs and trees. These classes represent plants of increasingly complex architecture^{2,7}. Insects that fed on plants of more than one growth form were assigned to the most structurally complex. Plant architecture has an important effect on herbivore species richness (see, for example refs 2 and 7) and may similarly influence parasitoids¹⁰.

(3) Number of plant species used by a host insect. These were counted using lists of all known British food plants (L. K. Ward, personal communication). Two variables were generated, including only plants native to Great Britain or both native and introduced plants. If herbivores are attacked by different species of parasitoids on different host plants (see, for example, ref. 11), the numbers of parasitoid species and host plants should be positively correlated.

(4) Degree of host-plant specialization. Phytophagous insects were classed according to the taxonomic relationships of their host plants. The insects were divided into five categories: strict monophages (feeding on one plant species); restricted to one plant genus; feeding on one plant family; feeding on two to four families; extreme polyphages (five or more plant families).

Fig. 1 Effects of four ecological variables on parasitoid species richness. Data are plotted as mean number of parasitoids per host on linear scales, but all associated statistics are based on ANOVA of logarithmically transformed richness. Vertical bars delimit significantly different means (Duncan's multiple range test, $P < 0.05$), and numbers are sample sizes. *a*, Range over which hosts have been studied ($F = 7.63$, $P = 0.0006$); *b*, host-plant architecture ($F = 8.17$, $P < 0.0001$); *c*, degree of specialization of host insects ($F = 5.65$, $P = 0.0002$ for native plant species, $F = 4.41$, $P = 0.0018$ for all plants); *d*, host insects' feeding niche ($F = 16.74$, $P < 0.0001$). See text for definitions of niches.



As with variable (3), separate analyses included only plants native to Britain or native plus introduced species. Variables (3) and (4) involve partially overlapping characteristics of the host insect and its range of food plants.

(5) Feeding niche. We recognized six herbivore feeding strategies based on their position on or in their host plants. In order of decreasing mobility and increasing concealment, these are: external herbivores; 'semi-concealed' species constructing simple structures, but capable of free movement when feeding (for example leaf folders); leaf miners; gall formers; species concealed within the plant, but without specialized structures associated with their presence (stem borers and those infesting wood or flowerheads) and root feeders. We expected both the degree of host mobility^{10,12} and concealment to influence parasitoid richness.

(6) Taxonomic isolation of the host insect. This was measured as the number of species in both the same genus and in the same family. We used the classifications of Kloet and Hincks¹³⁻¹⁷ with a few exceptions. We hypothesized that taxonomically isolated host insects (those with few relatives in the same genus) should recruit fewer parasitoids (see, for example, ref. 10).

(7) Host-insect development. The immature stages of insects with incomplete metamorphosis (exopterygotes) are more similar to one another than are those of insects with a pupal stage (endopterygotes). Therefore, the markedly different potential pupal niche should raise the average parasitoid burden on endopterygotes above that of exopterygotes.

Consistent with most studies of herbivores on plants^{1-5,7,18} extent studied/range has a positive effect on parasitoid richness (Fig. 1a). In our data most of this effect is undoubtedly due to passive sampling^{7,19,20}. Because many parasitoid complexes have not been intensively studied, area studied and sample sizes are related. More data are required before the non-biological effects of sample effort and the biologically more interesting effects of the geographic range of each host insect can be distinguished as determinants of parasitoid richness.

The positive influence of host-plant architecture is more meaningful (Fig. 1b). Available evidence suggests that this is a characteristic of insect communities at all trophic levels^{7,21,22}, although

the mechanisms may not be identical. For parasitoids, two processes may contribute to an effect of architecture. First, size *per se*; insects feeding on large, conspicuous plants are more likely to be encountered and colonized by parasitoids in evolutionary time. Second, as Askew has reasoned²², because there are increasing numbers of herbivore species and hence potential hosts as plants increase in size and complexity, polyphagous parasitoids may be attracted to large plants to use these additional host species.

Contrary to expectation, the degree of host-plant specialization by herbivores has an inconsistent impact on parasitoid richness (Fig. 1c). There are also no relationships between richness (log transformed) and the number of plants (log transformed) on which herbivores feed ($R^2 = 0.005$, $P = 0.237$ for native plants and $R^2 < 0.001$, $P = 0.998$ for all host plants). The parasitoids of British insects provide no support for the idea that herbivores feeding on many host plants should be exposed to more species of parasitoids, each searching a different subset of host plant species, than are herbivores restricted to a few host plants, despite the fact that many species of parasitoids are relatively host-plant specific²³⁻²⁶.

The effect of a herbivore's feeding strategy on parasitoid species richness, however, is unequivocal (Fig. 1d), and appears to be due to several interacting forces. First, as the mobility of hosts decreases from external and generally free-ranging feeders to much more stationary leaf miners, parasitoid richness increases twofold (although we question the magnitude of this difference because the parasitoids of many free-living herbivores may be undersampled). But once hosts have become relatively stationary (and the data more complete), parasitoid richness decreases with increasing host concealment and protection. Thus, gall formers have fewer parasitoids than leaf miners, whose presence in a mine is obvious to even casual searchers (such as entomologists), and whose mines offer minimum protection against attack. Even when galls have been discovered, the gall tissue makes the hosts themselves more difficult to locate and impedes parasitoid oviposition. In turn, gall formers have more parasitoids than species infesting stems, wood or flowerheads, which are both physically protected within plant tissues

Table 2 Parasitoid species richness per host species by order of host

Order	$\bar{x}$	Order	$\bar{x}$
Hemiptera	2.64*	Diptera	5.15†
Homoptera	4.02*	Hymenoptera	6.95†
Coleoptera	4.59*†	Lepidoptera	9.40‡

Means followed by same symbol not significantly different ($P > 0.05$, Duncan's multiple range test).

and generally do not provide visual clues to their presence. Lastly, root feeders are undoubtedly the most difficult insects to locate and generally have the fewest parasitoids.

The number of host-insect species in each genus shows a statistically significant, but very weak, relationship with parasitoid richness ($R^2 = 0.022$, $P = 0.012$). At the family level the relationship is not significant ($R^2 = 0.002$, $P = 0.424$). Thus, taxonomic isolation of hosts appears to have a minimal influence on parasitoid burdens.

The expectation that insects with incomplete metamorphosis should support fewer parasitoid species than those with complete metamorphosis is not well supported; Coleoptera do not have significantly more parasitoids than Hemiptera/Homoptera. Average parasitoid loads clearly differ among insect orders (Table 2), but may reflect differences due to factors other than development.

Drawing together our results, analysis of single factors suggests that some of the mechanisms contributing to the diversity of the second trophic level also operate at the higher trophic levels. Geographic range and measures of plant structural complexity are consistently correlated with herbivore richness^{7,21}, and we note similar effects here; effects of host-plant taxonomic isolation on herbivore richness are weaker and more variable^{5,7,27}, and taxonomic isolation of host insects also has a feeble effect on parasitoid richness. However, a factor missing from the second trophic level but acting on parasitoids is host feeding strategy. How and where a herbivore feeds is undoubtedly influenced by its natural enemies^{28,29}, and we expect the converse to be true²². Moreover, our data identify this as the most important factor contributing to parasitoid species richness.

In a step-wise multiple regression, host-insect feeding niche accounted for 20.5% of the variance in parasitoid richness, significantly more than any other variable. (Because our *a priori* arrangement of feeding niches was non-linear, a quadratic term was first fitted and the polynomial was allowed to enter the regression as a block.) Extent studied, explaining an additional 1.6%, significantly improved the model. Plant architecture and taxonomic isolation, although significant when analysed separately, failed to enter into the regression, as did all the remaining variables. That we can explain 22.1% of the variation in the species richness patterns of parasitoids is encouraging given the great limitations in the data, although because most variables are ordinal or nominal, interpretation of the regression can only be tentative. In general, however, similar models explain more of the variation in phytophagous insect species richness^{2,3,5,7}. It remains to be seen whether the diversity of higher trophic levels is inherently more stochastic than that of lower trophic levels, or whether further study and better data will improve the amount of variation explained by deterministic models.

We thank R. R. Askew, J. K. Waage and R. Law for their comments. L. K. Ward kindly provided host-plant data from the Phytophagous Insects Data Bank. This study was supported by a NSF Postdoctoral Fellowship in Environmental Biology to B.A.H.

- Claridge, M. F. & Wilson, M. R. *Ecol. Entomol.* 7, 19–30 (1982).
- Kennedy, C. E. J. & Southwood, T. R. E. *J. Anim. Ecol.* 53, 455–478 (1984).
- Godfray, H. C. J. *Ecol. Entomol.* 9, 163–168 (1984).
- Strong, D. R., Lawton, J. H. & Southwood, T. R. E. *Insects on Plants* (Blackwell, Oxford, 1984).
- Price, P. W. *Evolutionary Biology of Parasites* (Princeton University Press, 1980).
- Rey, J. R., McCoy, E. D. & Strong, D. R. Jr, *Am. Nat.* 117, 611–622 (1981).
- Askew, R. R. & Shaw, M. R. in *Insect Parasitoids* (eds Waage, J. & Greathead, D.) 225–264 (Academic, London, 1986).
- Ball, J. C. & Dahlsten, D. L. *Can. Ent.* 105, 1453–1464 (1973).
- Shaw, M. R. & Askew, R. R. in *A Dipterist's Handbook* (ed. Stubbs, A. & Chandler, P.) 164–171 (The Amateur Entomologist, vol. 15, 1979).
- Kloet, G. S. & Hincks, W. D. *A Checklist of British Insects. Part 1: Small Orders and Hemiptera* (Handbk. Ident. Br. Insects, Roy. Entomol. Soc. Lond., London, 1964).
- Kloet, G. S. & Hincks, W. D. *A Checklist of British Insects. Part 2: Lepidoptera* (Handbk. Ident. Br. Insects, Roy. Entomol. Soc. Lond., London, 1972).
- Kloet, G. S. & Hincks, W. D., *A Checklist of British Insects. Part 5: Diptera and Siphonaptera* (Handbk. Ident. Br. Insects, Roy. Entomol. Soc. Lond., London, 1976).
- Kloet, G. S. & Hincks, W. D., *A Checklist of British Insects. Part 3: Coleoptera and Strepsiptera* (Handbk. Ident. Br. Insects, Roy. Entomol. Soc. Lond., London, 1977).
- Kloet, G. S. & Hincks, W. D., *A Checklist of British Insects. Part 4: Hymenoptera* (Handbk. Ident. Br. Insects, Roy. Entomol. Soc. Lond., London, 1978).
- Stevens, G. C. *Am. Nat.* 128, 35–46 (1986).
- Connor, E. F. & McCoy, E. D. *Am. Nat.* 113, 791–833 (1979).
- Karban, R. & Ricklefs, R. E. *Ecology* 64, 636–641 (1983).
- Lawton, J. H. *A. Rev. Ent.* 28, 23–39 (1983).
- Askew, R. R. *J. Anim. Ecol.* 49, 817–829 (1980).
- DeBach, P. *Biological Control of Insect Pests and Weeds* (Chapman and Hall, London, 1964).
- Stary, P. *Biology of Aphid Parasites (Hymenoptera: Aphididae) With Respect to Integrated Control* (Junk, The Hague, 1970).
- Price, P. W. *et al. A. Rev. Ecol. Syst.* 11, 41–65 (1980).
- Vinson, S. B. in *Chemical Ecology of Insects* (eds Bell, W. J. & Carde, R. T.) 205–232 (Chapman and Hall, London, 1984).
- Cornell, H. V. *Ecology* 66, 1247–1260 (1985).
- Jefferies, M. J. & Lawton, J. H. *Biol. J. Linn. Soc.* 23, 269–286 (1986).
- Lawton, J. H. in *Insect Parasitoids* (eds Waage, J. & Greathead, D.) 265–287 (Academic, London, 1986).

Large proportion of marine planktonic ciliates found to contain functional chloroplasts

Diane K. Stoecker, Ann E. Michaels & Linda H. Davis

Biology Department, Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543, USA

Planktonic ciliates, most importantly tintinnids and oligotrichs (subclass Choreotrichia), are important components of the microplankton in coastal waters¹. At times, the biomass of ciliates is equivalent to the biomass of larger zooplankters². The role of ciliates as grazers has been recognized (see ref. 3 for review), but their role as chloroplast-bearing cells has not been widely appreciated⁴. Many planktonic oligotrichs are pigmented because they sequester chloroplasts derived from a variety of chromophytic algae^{5–7}. We report the first quantitative data on the frequency of chloroplast retention among planktonic oligotrichs and tintinnids. In surface waters during the spring and summer, approximately 42% of the planktonic ciliates have chloroplasts. During the autumn and winter, chloroplast-retaining species are less abundant, but still constitute at least 10% of the combined tintinnid and oligotrich fauna. Water samples containing large numbers of pigmented ciliates have a high chlorophyll content⁸ and fix carbon⁹. Although it has been speculated that sequestered chloroplasts in ciliates are functional^{5–7}, we have demonstrated photosynthesis by chloroplast-containing ciliates experimentally. The trophic position of ciliates in planktonic food webs needs to be reconsidered.

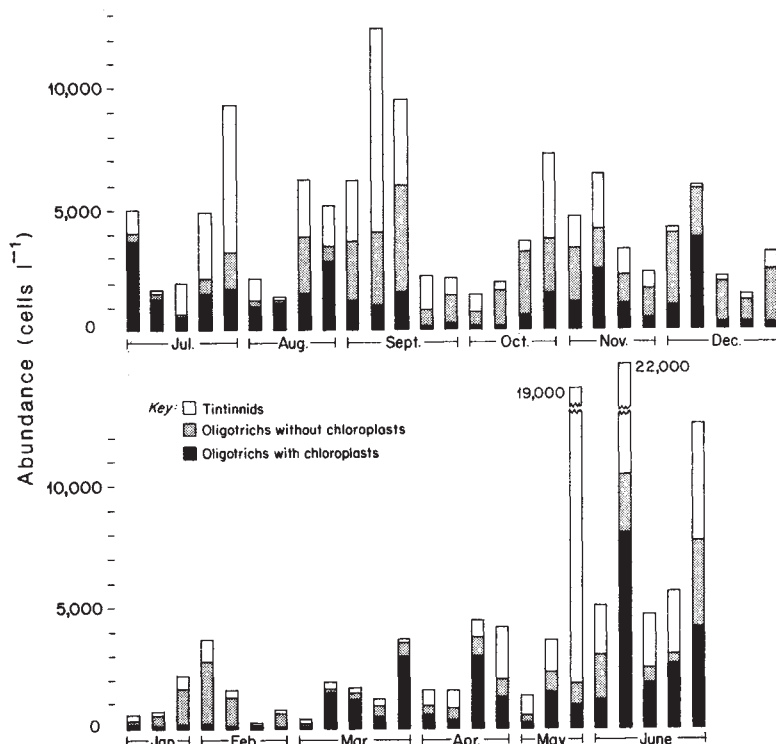
Surface water samples were collected from the Oceanographic Institution dock in Great Harbor, Woods Hole, Massachusetts and the planktonic ciliates were enumerated (Fig. 1). This location is well mixed by tidal currents, and thus stratification of the water column is minimal, and samples could be collected all round the year.

In each sample, we observed between one and eight distinct types of oligotrichs with chloroplasts. These included *Laboea strobila*, *Tontonia* sp., and several *Strombidium* spp. In our samples, none of the tintinnids sequestered chloroplasts. With transmitted light, the ciliates with chloroplasts were greenish,

Received 5 December 1986; accepted 19 February 1987.

- Opler, P. A. *Am. Nat.* 62, 67–73 (1974).
- Lawton, J. H. & Schroeder, D. *Nature* 265, 137–140 (1977).
- Lawton, J. H. & Price, P. W. *J. Anim. Ecol.* 48, 619–637 (1979).

Fig. 1 Estimated abundances of tintinnids (open bars) and oligotrichs, with (filled bars) and without (hatched bars) chloroplasts, in surface water samples from Great Harbor, Woods Hole, Massachusetts (41° 31' N, 70° 41' W; depth 17 m). These groups represent over 95% of the ciliates present. Sampling began in July 1985 and continued until June 1986. Samples were collected weekly and were fixed in 2% buffered formaldehyde and stored at 4°C in the dark for less than 7 days before being examined. Ciliates in 100-ml subsamples were counted in settling chambers with an inverted microscope. Chloroplast-retaining species can be unambiguously distinguished from species without chloroplasts using epifluorescence microscopy (a Zeiss inverted microscope equipped with a 50 W Mercury lamp, a BP 450–490 nm excitation filter, a LP 520 nm barrier filter, and a FT 510 chromatic beam splitter). Epifluorescence due to retained chloroplasts occurs throughout the cell but is concentrated near the periphery; the individual bodies are small and intensely epifluorescent. Losses due to fixation¹⁰ probably resulted in a 10–20% underestimate of the oligotrich population.



but the shade and intensity of coloration varied. Almost all the chloroplasts in the ciliates displayed red epifluorescence, which indicates that they contained chlorophyll *a* but not phycoerythrin (if phycoerythrin is present, the epifluorescence is yellow-orange). Over the year, an average of 45% (s.d. 27%) of the oligotrich cells had chloroplasts (Fig. 1). In some of the summer samples, over 90% of the oligotrichs had chloroplasts. An average of 31% (s.d. 22%) of the ciliates (tintinnids and oligotrichs) were chloroplast-retaining oligotrichs; this estimate is probably low due to differential loss of oligotrichs during fixation¹⁰.

In Great Harbor, the average density of ciliates with chloroplasts was 1,167 cells l⁻¹ (s.d. 1,338). From the cell counts and calculated cell volumes, the carbon content of the chloroplast-retaining ciliates was estimated¹. In June, when oligotrichs with chloroplasts were most abundant (Fig. 1), their average density was 3,626 cells l⁻¹ (s.d. 2,766) and their estimated biomass was 6.5 µg C (micrograms carbon) l⁻¹ (s.d. 3.3). Over the year, the average biomass of chloroplast-retaining ciliates was 2.0 µg C l⁻¹ (s.d. 2.2). We do not have biomass estimates for the total ciliate fauna in Great Harbor, but biomass estimates for the ciliates in other coastal waters range from 0.9 to 9.0 µg C l⁻¹. Oligotrichs with chloroplasts make an important contribution to the total biomass of ciliates.

During July 1986, we collected surface water samples from Vineyard Sound, Buzzards Bay and Nantucket Sound to determine if chloroplast-retaining ciliates were generally abundant (Table 1). We found chloroplast-retaining species in all three bodies of water and found that an average of 69% of the oligotrichs had chloroplasts; these chloroplast-containing cells accounted for 50% of the combined tintinnid and oligotrich fauna.

We investigated the vertical distribution of planktonic ciliates by sampling at 3-m depth intervals with a Niskin bottle at station 'K' (Table 1). Because tidal currents are weaker at this station than in any of Great Harbor, Vineyard Sound, or most of Nantucket Sound¹¹, any changes in ciliate abundance with depth would be at their most pronounced there. Samples were taken at midday on two days. The average water column density of oligotrichs with chloroplasts was slightly higher than the surface density, but the relative contribution of chloroplast-retaining

species to the ciliate assemblage was 25–32% higher at the surface than in the water column as a whole. Thus, surface water samples appear to overestimate the frequency of chloroplast retention but slightly underestimate the abundance of chloroplast-retaining ciliates in the water column.

From Great Harbor, we isolated three chloroplast-retaining oligotrichs, *Laboea strobila* and two undescribed species of *Strombidium*. The ciliates were grown in a seawater medium¹² on a 14:10 h light:dark cycle at 15°C. The microalgae, *Isochrysis galbana* and *Chroomonas salina*, were provided as food. Algal food is required to maintain growth of these ciliates. All three

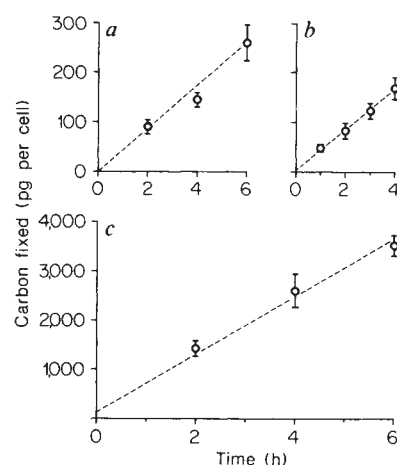


Fig. 2 Photosynthesis (uptake of ¹⁴C in the light minus uptake in the dark) by three oligotrichs that sequester chloroplasts, *Strombidium* sp. A (a), *Strombidium* sp. (b), and *Laboea strobila* (c). Vertical bars, standard deviations about the mean. The ciliates were separated from their algal food approximately 12 h before the experiments, which were conducted in the absence of algae. The ciliates were incubated in medium containing an estimated 0.5 µCi ml⁻¹ of ¹⁴C-bicarbonate¹³ in the light and in the dark at 15°C. At each sampling, 20 *Strombidium* or 15 *Laboea* were picked, with a drawn-out Pasteur pipette, into each of three replicate scintillation vials, where they were prepared for liquid scintillation counting¹⁴.

Table 1 Survey of tintinnids and oligotrichs, with and without chloroplasts, in surface waters of Vineyard Sound, Buzzards Bay and Nantucket Sound

Location	Station	Coordinates		Oligotrichs with chloroplasts	Abundance (cells l ⁻¹)		Oligotrichs with chloroplasts (% total)	Biomass of oligotrichs with chloroplasts (µg carbon l ⁻¹)
		N	W		Oligotrichs without chloroplasts	Tintinnids		
Vineyard Sound 6 July	A	41°30'0"	70°41'10"	3179	836	1056	63	2.6
	B	41°28'10"	70°44'40"	2398	1199	924	53	3.1
	C	41°26'15"	70°47'0"	2288	1342	957	50	3.1
	mean			2622	1126	979	55	3.0
	s.d.			397	213	56	5	0.2
Buzzards Bay 6 July	D	41°28'0"	70°48'50"	4213	1232	836	67	5.4
	E	41°29'40"	70°46'25"	2849	1892	4312	31	3.6
	F	41°31'30"	70°44'30"	2332	1353	3498	32	3.9
	mean			3131	1492	2882	44	4.3
	s.d.			793	287	1484	17	0.8
Nantucket Sound 15 July	G	41°28'20"	70°29'0"	1925	748	704	57	1.7
	H	41°27'40"	70°23'35"	1353	693	682	50	1.7
	I	41°27'50"	70°17'30"	957	825	1386	30	1.1
	J	41°26'35"	70°11'0"	2728	594	1122	61	2.0
	mean			1741	715	974	50	1.6
	s.d.			666	84	296	12	0.3
	1 July	K (0 m)	41°20'0"	2706	0	88	97	0.2
		(3 m)		8206	176	5918	57	14.7
		(6 m)		781	198	451	55	0.9
		(9 m)		1364	165	1276	49	1.0
Nantucket Sound 1 July	K (0 m)	41°20'0"	70°10'0"	1023	88	297	73	1.3
		(3 m)		1859	264	561	69	2.6
		(6 m)		1166	132	3003	27	1.3
		(9 m)		528	99	1529	24	0.7
	mean			1144	146	1348	48	1.5
	s.d.			476	70	1060	23	0.7

Depth at stations: Vineyard Sound, 18–27 m; Buzzards Bay, 13 m; Nantucket Sound, 11–20 m. At one station, 'K' in Nantucket Sound, samples were taken at 3-m depth intervals at midday. Sample fixation, storage, and examination were the same as for the Woods Hole samples. None of the tintinnids contained chloroplasts. Losses due to fixation⁸ probably resulted in a 10–20% underestimate of the oligotrich fauna.

species contain both digestive vacuoles and intact chloroplasts derived from both algal species. The ciliates maintain chloroplasts even if they are starved or kept in the dark; however the chlorophyll content and number and size of chloroplasts vary with growth conditions (D.K.S., A.E.M. & L.H.P., unpublished data).

We grew the ciliates and measured their photosynthetic rate at a photosynthetically available radiation (PAR) flux of 150–200 µE m⁻² s⁻¹. Photosynthesis was measured by ¹⁴C uptake (Fig. 2). Because sequestered chloroplasts may preferentially take up metabolic carbon dioxide produced in the ciliate, these measurements are minimum estimates of the rate of photosynthesis by the chloroplasts. At a light flux of 150–200 µE m⁻² s⁻¹, the uptake of ¹⁴C was linear over 6 h and appreciable photosynthesis occurred in all three species (Fig. 2). *Laboea strobila*, with a cell volume of approximately 1.1 × 10⁵ µm³ and a chlorophyll *a* content of 187 pg (s.d. 58), had the highest photosynthetic rate, 586 pg carbon per cell per hour (3.1 pg carbon per pg chlorophyll *a* per h). *Strombidium* sp. A, with a cell volume of 7.6 × 10⁴ and a chlorophyll *a* content of 57 pg (s.d. 10), had a rate of 42 pg carbon cell⁻¹ h⁻¹ (0.7 pg carbon per pg chlorophyll *a* per h). *Strombidium* sp. B, with a cell volume of only 1.4 × 10⁴ µm³ and a chlorophyll *a* content of 20 (s.d. 1), had a photosynthetic rate of 41 pg carbon cell⁻¹ h⁻¹ (2.1 pg carbon per pg chlorophyll *a* per h). The rate of photosynthesis is not strictly proportional to cell size or to chlorophyll content. Some oligotrichs which retain chloroplasts appear to be more autotrophic than others.

Planktonic ciliates which retain chloroplasts are an important, year-round component of the microplankton. Chloroplast-retaining oligotrichs have been reported from far-ranging environments including the Pacific and Atlantic oceans, and the Mediterranean Sea^{5–7}. We have observed chloroplast-retaining oligotrichs in samples collected on George's Bank, in nearshore

waters of the northern Gulf of Mexico, and in estuaries on Cape Cod as well as from local bays and sounds. Chloroplast retention by planktonic ciliates appears to be a widespread phenomenon. Our data demonstrate that at least three common chloroplast-retaining species are mixotrophic, acting both as grazers on phytoplankton and as photosynthetic autotrophs by means of their sequestered chloroplasts. Of these three, *Laboea strobila* is the only one which is easily classified to species in natural samples. It accounted for 46% of the biomass of chloroplast-retaining oligotrichs in Woods Hole. Based on these results, we suggest that oligotrichous ciliates are important both as producers and consumers in marine planktonic communities. Because of mixotrophy, species with chloroplasts may be extremely efficient producers of animal biomass in the oceans⁷.

Supported by grants from the NSF. Preliminary ¹⁴C experiments were done in collaboration with J. C. Goldman. We thank D. J. S. Montagnes and E. B. Small for identifying *Laboea strobila*, and D. A. Caron and J. B. Waterbury for their comments.

Received 1 December 1986; accepted 18 February 1987.

1. Beers, J. R., Reid, F. M. H. & Stewart, G. L. *Mar. Biol.* **60**, 209–226 (1980).
2. Smetacek, V. *Mar. Biol.* **63**, 1–11 (1981).
3. Verity, P. G. *Mar. Ecol. Prog. Ser.* **29**, 105–115 (1986).
4. Elbrachter, M. in *Flows of Energy and Materials in Marine Ecosystems* (ed. Fasham, M. J. R.) 191–222 (Plenum, New York, 1984).
5. Blackburn, D. J., Taylor, F. J. R. & Blackburn, J. *J. Protozoology* **20**, 286–288 (1973).
6. Laval-Peuto, M. & Febvre, M. *Biosystems* **19**, 137–158 (1986).
7. McManus, G. B. & Fuhrman, J. A. *J. Plank. Res.* **8**, 317–327 (1986).
8. Holm-Hansen, O., Taylor, F. J. R. & Barsdate, R. J. *Mar. Biol.* **7**, 37–46 (1970).
9. Burkholder, P. R., Burkholder, L. M. & Almodovar, L. R. *Bull. Mar. Sci. Gulf Caribb.* **17**, 1–15 (1967).
10. Gifford, D. J. *Mar. Ecol. Prog. Ser.* **23**, 257–267 (1985).
11. *Eldridge Tide and Pilot Book*, 1986 (Robert Eldridge White, Boston, 1985).
12. Stoecker, D., Guillard, R. R. L. & Kavee, R. M. *Biol. Bull. mar. biol. Lab., Woods Hole* **160**, 136–145 (1981).
13. Rivkin, R. B. & Seliger, H. H. *Limnol. Oceanogr.* **26**, 780–785 (1981).
14. Li, W. K. W. & Goldman, J. C. *Microbiol. Ecol.* **7**, 113–121 (1981).

Peripheral modulation of mechanosensitivity in primary afferent neurons

Valerie M. Pasztor* & Brian M. H. Bush

Department of Physiology, University of Bristol, Bristol BS8 1TD, UK

* Biology Department, McGill University, 1205 Dr Penfield Avenue, Montreal H3A 1B1, Canada

Considerable attention has centred recently on the changes in neuron excitability¹ and synaptic efficacy² caused by certain biogenic amines and neuropeptides. These neuromodulators act at a wide variety of both central and peripheral targets, and bring about diverse biological results. In sensory pathways, modulation occurs at central input synapses of the primary afferents^{3,4} and at peripheral terminals of efferents⁵⁻⁷. This study was undertaken to look at non-synaptic modulation of membrane potentials in peripheral sensory endings of identifiable receptors. Using intracellular recording from the three primary afferent fibres of a recently described simple crustacean stretch receptor, which lacks centrifugal control, we observed *in vitro* modulation of the sensory response by three neuroactive substances known to be present *in vivo*. Two neuroamines, serotonin and octopamine, depressed receptor potentials and impulse discharge whereas the pentapeptide proctolin enhanced both these components of the sensory response. The peripheral sensory modulation reported here for a lobster mechano-receptor may occur in many animal groups and sensory systems.

The stretch receptor used was the lobster oval organ⁸, which provides feedback from the ventilatory appendage to the central nervous system (Fig. 1a and b). It has distinct advantages for a study of this kind because it has three large sensory neurons (X, Y and Z), with central cell bodies and afferent fibres large enough to permit intracellular recording (Fig. 1c). Their unusually large length constants (>10 mm) result in the receptor potentials being conducted with relatively little decrement^{9,10} from peripheral transducer sites to central neuropile. Unlike other crustacean mechano-receptors capable of long distance conduction of graded signals¹¹, however, oval organ afferents generate normal frequency modulated patterns of nerve impulses. Thus, when recording responses to mechanical stimulation from exposed, immobilized sensory afferents, both receptor potentials and regenerative impulses may be observed. Other advantages derive from the simplicity of the oval organ structure. The sensory dendrites are supported by an array of connective tissue strands (Fig. 1c) with neither receptor muscle nor any other means of efferent control, so peripheral synaptic modulation plays no role in this system.

Figure 2 shows examples of responses of the three fibres to trapezoidal pull stimuli. One-minute inter-stimulus periods ensured that the sensory units regained the fully rested state. We have shown previously¹² that fibres X, Y and Z, though anatomically similar in their peripheral arborization patterns, have characteristically distinctive afferent responses: X fibres develop the largest receptor potential amplitudes, Z fibres have the highest thresholds but attain the highest firing frequencies, and Y fibres show intermediate properties. In most preparations X fibres generated only a brief burst of impulses limited to the positive dynamic phase of a ramp-and-hold (trapezoidal) stimulus (Fig. 2a). However, some fibres, identified as X on the basis of diameter and large receptor potential, showed an extended firing pattern which more closely resembled that of a Y fibre (Fig. 2c). Because differences in responses to neuromodulators were also observed, we present data from 'normal' X fibres (type X₁) separately from those showing the unusual spiking pattern (type X₂).

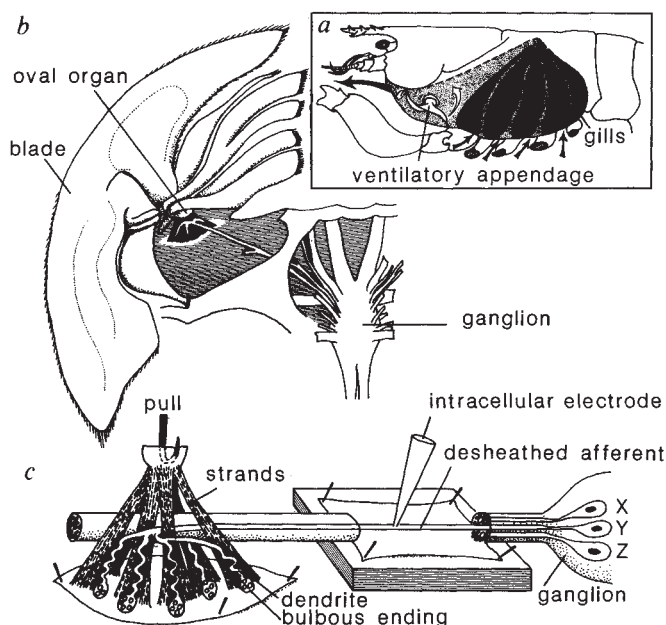


Fig. 1 The location of the oval organ in the ventilatory appendage and the isolated preparation used for intracellular recording. *a*, Diagram of the lobster branchial chamber showing the water currents (black arrows) driven over the gills by the beating of the ventilatory appendage (white arrows). The oval organ is situated at the fulcrum of the twisting motion of the blade. *b*, The oval organ *in situ* (not to scale) showing sensory dendrites arborizing through a cone of connective tissue strands, spanning a gap at the base of the blade. *c*, The isolated preparation. Only one of the three afferent fibres is shown in its entirety. A hook attaches a fragment of exoskeleton at the apex of the cone of connective tissue strands to an electromagnetic puller. The cuticle at the base of the cone is securely pinned to the Sylgard lining of the bath. Upward pull on the strands stretches the dendrites. The nerve sheath is pinned back to Sylgard platform with cactus spines to immobilize the exposed sensory afferents during mechanical stimulation. Afferent diameters: X, 41 μm; Y, 32 μm; Z, 22 μm. Distance from confluence of sensory dendrites to ganglion, 11 mm; to recording site, 5 mm. The preparation was superfused with Tris-buffered lobster saline¹⁰. Test neuroactive substances were added to the bath over a concentration range of 10⁻¹¹ M to 10⁻⁵ M. Temperature of bath and perfusates was strictly maintained at 15 °C.

Proctolin enhanced the sensory responses in X₂, Y and Z fibres. It increased numbers of spikes and amplitudes of graded potentials, as compared with the controls (Table 1). Action potentials and their negative-going afterpotentials obscure the actual magnitude of the receptor potential, particularly during the dynamic phase. However, measurements for Table 1 taken at the end of the static phase plateau, after cessation of spiking, show clear increases in receptor potential amplitude with proctolin application (for example Fig. 2b, fibre Y). This was particularly clear with responses which were sub-threshold or just supra-threshold for spiking (Fig. 3). In contrast to X₂ fibres, X₁ fibres were unaffected by proctolin (Fig. 2a). Serotonin reduced sensory responsiveness in fibres Y and Z: the receptor potential declined and, in the presence of 2.5 × 10⁻⁶ M serotonin, failed to reach spiking threshold (Figs 2c, 3b and c; Table 1). There is no evidence for any change either in spiking threshold or in the ability to generate spikes, because by increasing the pull amplitude, thus evoking a larger receptor potential, normal spiking was restored. Unlike Y and Z fibres, both types of X fibre were resistant to serotonin. The effects of octopamine were neither as consistent as those of proctolin and serotonin, nor as pronounced with this stimulus regime. Octopamine depressed responses in 40% of the trials, and either had no effect or produced a weak enhancement in the remainder.

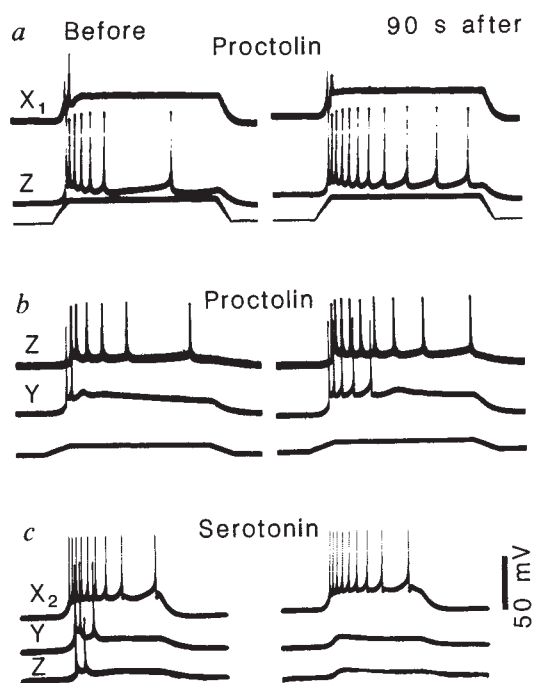


Fig. 2 Effects of 2.5×10^{-6} M proctolin (*a* and *b*) and serotonin (*c*) on spiking patterns of oval organ afferents. Intracellular recordings made simultaneously from two or three afferents. Standard one-second pull stimuli were delivered at one-minute intervals (bottom traces in *a* and *b*; not included in *c*). *a*, Spiking in fibre X_1 rose transiently from 7 to 11 spikes per response, whereas fibre X_1 showed no change in either number of spikes or amplitude of graded potential. *b*, In both fibres Y and Z , proctolin enhanced spiking. *c*, Receptor potentials in Y and Z , were depressed by serotonin to levels below the spiking threshold. Fibre X_2 showed no change.

In vivo the oval organ is subjected to continuous sinusoidal stretch stimulation as the ventilatory appendage beats up and down, pumping water through the branchial system. To simulate this more natural condition, continuous trains of 500-ms pulls were delivered at 1 Hz. Under this stimulus regime, the responses of the primary afferents show 'habituation'¹³. Receptor potential amplitudes, and resulting bursts of spikes, decline progressively during a train of repetitive pulls in normal saline (Fig. 4*a*), and after 10–15 stimuli receptor potentials may fail to reach spiking threshold. Generally the responses stabilize to a pattern of supra-threshold (spiking) and sub-threshold (non-spiking) responses. The proportion of spiking responses can be adjusted to a desired level by manipulating the amplitude and frequency of the pull stimuli. This provides a sensitive test for modulation and was used to determine threshold concentration and dose dependency.

When proctolin was applied to a preparation fully habituated to 1-Hz stimulation, the proportion of responses in fibres Y and Z in which the receptor potential reached spiking threshold rose in a dose dependent manner (Fig. 4*b*). The threshold for a just-detectable enhancement was 2.5×10^{-10} M. Fibre X_2 in this example showed no change in spiking, although the receptor potential was slightly enhanced and showed less decline during the static plateau. Both serotonin (Fig. 4*c*) and octopamine reduced the number of responses producing spikes in fibres Y and Z , but had no effect on X fibres. Their effects were similar in magnitude and time course, with octopamine having a lower threshold concentration (2.5×10^{-11} M) than serotonin (2.5×10^{-10} M).

Proctolin enhancement was transient at all concentrations, recovery usually starting within 3–6 minutes even when exposure to the substance was prolonged (Fig. 3*a*). By contrast, the

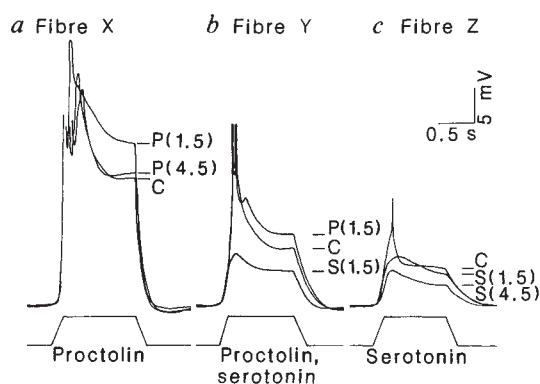


Fig. 3 Three representative examples of modulation of receptor potential by 2.5×10^{-6} M proctolin and serotonin. Standard one-second pulls given at one-minute intervals as in Fig. 1. Chart recordings retraced and superimposed to show changes in static plateau amplitude (spikes truncated). C, Control response just before addition of each modulator; P, S, responses with proctolin or serotonin in the bath, recorded 1.5 or 4.5 min (in parentheses) after addition in each case. *a*, Fibre X_2 receptor potential is transiently enhanced by proctolin: the maximum increase was recorded 1.5 min after adding proctolin to the bath, and recovery to the control level was almost complete by 4.5 min after proctolin addition. *b*, Fibre Y response is enhanced by proctolin and reduced by serotonin (two tests on the same preparation separated by one hour washout). *c*, Fibre Z response shows progressive reduction in presence of serotonin (with no subsequent recovery).

Table 1 Changes in the responses of oval organ afferents to pull stimuli induced by proctolin and serotonin

Fibre	Change in amplitude of static plateau* (%)			Change in number of spikes† (%)		
	Mean	s.e.m.	No. of observations	Mean	s.e.m.	No. of observations
Proctolin (2.5×10^{-6} M)						
X_1	0	0	4	0	0	5
X_2	+21.5	4.2	3	+16	13.5	3
Y	+26.8	1.8	4	+58.8	25.8	9
Z	+29.8	7.7	4	+62.9	24.3	7
Serotonin (2.5×10^{-6} M)						
X_1	0	0	2	0	0	2
X_2	0	0	1	0	0	2
Y	-43.2	13.0	5	-85	15.2	5
Z	-42.5	0.5	2	-100	0	4

* Measurements taken only from responses showing clear static plateaux undistorted by spike afterpotentials as exemplified by Fig. 3.

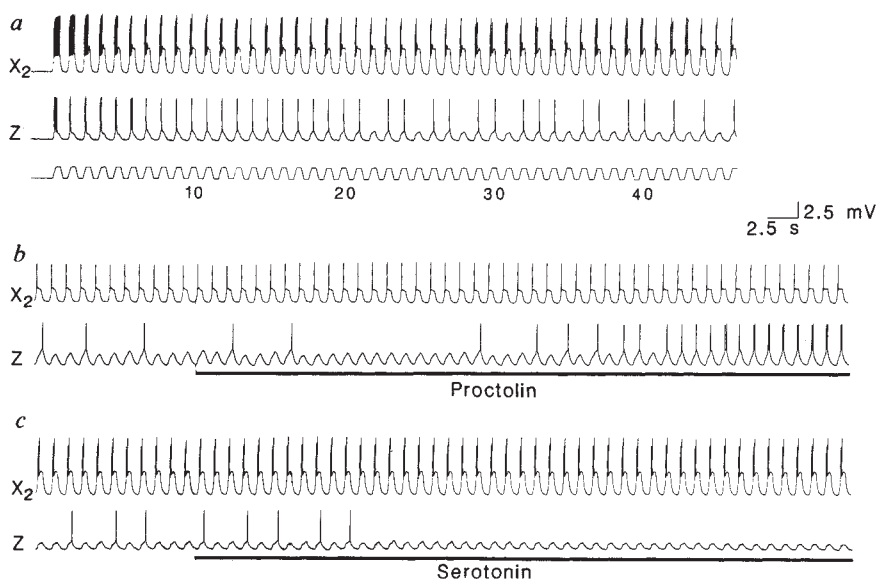
† Counts of impulses in responses to standard trapezoidal pulls as shown in Fig. 2.

response depression by serotonin and octopamine was long lasting at the highest concentrations (over 10^{-8} M or 10^{-5} M, respectively), and was frequently irreversible in spite of assiduous washout.

The two amines serotonin and octopamine, have been detected in the haemolymph of lobsters at concentrations of $1-5 \times 10^{-8}$ M (ref. 14), and act as circulating hormones with a variety of effects on target cells, notably neuromuscular junctions¹⁵. The threshold values of 10^{-10} – 10^{-11} M that we determined *in vitro* for sensory modulation are well within the available circulating levels, suggesting that neurohormonal gain-control of these proprioceptors *in vivo* is indeed a possibility.

Proctolin is also present in lobsters. However, assays of haemolymph failed to show significant concentrations of free proctolin, and injections of exogenous proctolin disappeared from the circulation with a half life of 6–10 minutes¹⁶. Possibly

Fig. 4 Effects of 2.5×10^{-8} M proctolin and serotonin on spiking and receptor potentials of X_2 and Z fibres following habituation to 1 Hz stimulation. For each of the three extracts, recorded from a single preparation (with a one hour interval between *b* and *c*), pull amplitude was adjusted to give an appropriate ratio of spiking to non-spiking responses in Z. *a*, Development of habituation at the beginning of repetitive stimulation in normal saline, showing progressive decline in numbers of spikes per response. After 22 pulls, receptor potential amplitude in fibre Z was close to the firing threshold; the response pattern then stabilized to an alternation of spiking and non-spiking responses. Habituation in fibre X_2 was less pronounced than in Z, and X_2 receptor potential amplitude never fell below spiking threshold. *b*, Proctolin promoted receptor potential growth in fibre Z to a supra-threshold magnitude which then supported two spikes per response. Spiking in fibre X_2 in this example was unchanged despite a 20% increase in the static phase of receptor potential. *c*, Fibre X_2 was unresponsive to serotonin, whereas in fibre Z receptor potential amplitude declined and spiking fell to zero.



proctolin release sites are strategically located close to the target tissues. Recent immuno-cytochemical studies (K.K. Siwicki and V.M.P., unpublished) have shown proctolin-like immunoreactivity not only in the three afferent fibres of the oval organ, but also in their extensive peripheral dendritic arborizations. This, together with earlier observations of dense cored vesicles in oval organ sensory endings⁸, suggests that the mechanoreceptive dendrites might serve a dual role as both transducer and neurosecretory release site. Whether proctolin is indeed released from the dendritic terminals during normal ventilatory activity *in vivo*, and if so, whether it has a purely local feedback effect on the sensory signal itself, or a more general modulatory role following dispersion through the haemolymph, remains to be investigated.

The transient nature of the proctolin effects observed with the oval organ contrasts markedly with the long-lasting effects commonly reported for neurohormones^{3,5,6}. However, hormonal actions in general form a continuum, ranging in duration from minutes to days. In this system proctolin provides positive gain in a proprioceptive feedback loop, so it would seem appropriate that it has only a short half life in the general circulation and a similarly brief action upon the transducer membrane.

This work was supported by research grants from the NSERC (Canada) to V.M.P. and from the SERC (UK) to B.M.H.B. We thank Malcolm Burrows, Robert Elson, Stewart Head, Mike Levine and Mel Robertson for helpful comments on the manuscript.

Received 27 November 1986; accepted 13 February 1987.

1. Aréchiga, H., García, U. & Rodríguez-Sosa, L. in *Model Neural Networks and Behavior* (ed. Selverston, A. I.) 361-379 (Plenum, New York, 1985).
2. Illes, P. *Neuroscience* 17, 909-928 (1986).
3. Pollock, J. D., Bernier, L. & Camardo, J. S. *J. Neurosci.* 5, 1862-1871 (1985).
4. Glanzman, D. L. & Krasne, F. B. *J. Neurosci.* 3, 2263-2269 (1983).
5. Kass, L. & Barlow, R. B. *J. Neurosci.* 4, 908-917 (1984).
6. Mancillas, J. R. & Selverston, A. I. *J. Neurosci.* 4, 847-859 (1984).
7. Wu, R. & Farley, J. *Soc. Neurosci. Abstr.* 10, 620 (1984).
8. Pasztor, V. M. *Zoomorphologie* 193, 171-191 (1979).
9. Pasztor, V. M. & Bush, B. M. H. *Science* 215, 1635-1637 (1982).
10. Pasztor, V. M. & Bush, B. M. H. *J. exp. Biol.* 107, 431-449 (1983).
11. Bush, B. M. H. in *Neurons without Impulses, Soc. Exp. Biol. Seminar Ser.* Vol. 6 (eds Roberts, A. & Bush, B. M. H.) 147-176 (Cambridge University Press, 1981).
12. Bush, B. M. H. & Pasztor, V. M. *J. exp. Biol.* 107, 451-464 (1983).
13. Pasztor, V. M. & Bush, B. M. H. *J. exp. Biol.* 107, 465-472 (1983).
14. Livingstone, M. S., Schaeffer, S. F. & Kravitz, E. A. *J. Neurobiol.* 12, 27-54 (1981).
15. Kravitz, E. A. *et al.* in *Model Neural Networks and Behavior* (ed. Selverston, A. I.) 339-360 (Plenum, New York, 1985).
16. Schwarz, T. L., Lee, G. M.-H., Siwicki, K. K., Standaert, D. G. & Kravitz, E. A. *J. Neurosci.* 4, 1300-1311 (1984).

Regulation of cytolytic T-lymphocyte generation by B-cell stimulatory factor

Michael B. Widmer & Kenneth H. Grabstein

Immunex Corporation, 51 University Street, Seattle, Washington 98101, USA

The growth and differentiation of cytolytic T lymphocytes (CTL) is regulated by soluble growth hormones, of which interleukin-2 (IL-2) is considered to be of prime importance. Here we report that the lymphokine B-cell stimulatory factor (BSF-1 or interleukin-4) also has profound effects on the generation of these functionally active T cells. In particular, BSF-1 acts as a potent helper factor for the generation of CTL in primary mixed leukocyte culture (MLC) and induces cytolytic activity in *in vitro* primed, MLC memory populations. Direct comparison of purified recombinant BSF-1 and IL-2 reveals BSF-1 to be the more potent CTL helper factor in primary MLC. Interestingly, the two lymphokines differed in that IL-2, but not BSF-1, induced a lytic population in cultures of unprimed cells without an overt antigenic stimulus. Collectively, our data provide a direct demonstration of a heretofore undefined mechanism by which CTL activation and amplification can occur.

BSF-1, characterized by its effects on B cells¹, has been shown recently to stimulate proliferation of certain factor-dependent, non-B lineage cell lines that are normally responsive to IL-2 or to the myeloid growth factor interleukin-3^{2,3}. To determine whether BSF-1 also affects primary T-cell populations, particularly with regard to the generation of functionally active T cells, we examined its effects on the generation of CTL in MLCs established with a suboptimal concentration of C57BL/6 splenic responding cells and allogeneic, irradiated, DBA/2 splenic stimulating cells. Five days after initiation of the culture, lytic activity against ⁵¹Cr-labelled P815 (DBA/2 origin) tumour target cells was measured. As shown in Fig. 1, cultures supplemented at initiation with 2 ng ml⁻¹ of homogeneous natural BSF-1 showed ~50-fold greater cytolytic activity per cell, and 100-fold more activity per culture, than control cultures in which exogenous BSF-1 was not present. Cultures supplemented with 10 ng ml⁻¹ recombinant IL-2 (rIL-2), as expected, also showed

Table 1 Enhancement by BSF-1 of alloreactivity of purified splenic CD8⁺ cells

T-cell depleted stimulating cells	Culture supplement	Microcultures ³ H-thymidine incorporation (c.p.m.)	Macrocultures	
			Recovery (%)	LU per culture
C57BL/6	Medium	418 ± 336	ND	ND
	rIL-2	11,467 ± 985	ND	ND
	rBSF-1	3,215 ± 528	ND	ND
DBA/2	Medium	5,552 ± 1,088	29	6
	rIL-2	21,977 ± 2,414	78	192
	rBSF-1	35,801 ± 1,716	98	714

ND, not determined; LU, lytic units (see Fig. 1 legend). T lymphocytes from C57BL/6 spleen were enriched by passage over nylon wool columns¹⁶. CD8⁺ cells were purified from nylon nonadherent cells by a combination of anti-CD4 monoclonal antibody GK1.5 (ref. 17) and complement depletion followed by positive selection on a Petri dish coated with anti-Lyt2.2 monoclonal antibody (83-12-5 from Dr Jeffrey Bluestone of NIH). Purified cells were analysed by flow microfluorimetry for CD4 expression by staining with GK1.5 followed by fluorescein isothiocyanate (FITC)-anti-rat κ (Becton Dickinson) or for CD8 expression by staining with 83-12-5 followed by FITC-goat F(ab')₂-anti-mouse Immunoglobulin M (IgM) (Tago). This analysis indicated that the purified cells contained no detectable CD4⁺ cells and at least 99.3% CD8⁺ cells. Stimulating cells were depleted of T cells by treatment with an antibody cocktail containing rabbit anti-mouse thymocyte antiserum (bone marrow and liver absorbed), GK1.5, and T24-31.7 anti-Thy1 monoclonal antibody¹⁸ plus rabbit complement. These cells were completely non-responsive to the T-cell mitogen concanavalin A (Con A), but retained full responsiveness to the B-cell mitogen lipopolysaccharide (LPS). To determine the effects of BSF-1 on cellular proliferation, MLC were established in 96-well microtitre plates with 1×10^5 CD8⁺ cells plus 1×10^6 T-cell-depleted, irradiated (2500R) C57BL/6 or DBA/2 splenic stimulating cells and 10 ng ml^{-1} purified recombinant BSF-1 or IL-2. Incorporation of ³H-dT was determined on day 4 after a 20-h pulse. MLCs for CTL generation were established in 16-mm culture wells (macrocultures) with 3.7×10^6 C57BL/6 CD8⁺ cells, 5×10^6 T-depleted, irradiated DBA/2 cells and 10 ng ml^{-1} of rIL-2 or rBSF-1. CTL activity against P815 targets was determined after four days as in the Fig. 1 legend. Spontaneous release was 129 c.p.m. and maximum release, 1,349 c.p.m. BSF-1 was cloned from a cDNA library made from size-fractionated mRNA from PMA-stimulated EL4 thymoma cells, and rBSF-1 was expressed in yeast using the alcohol dehydrogenase 2 (*ADH2*) promoter and the α -factor gene leader sequence to direct synthesis and secretion. The rBSF-1 was purified to homogeneity from yeast supernatant by reverse phase HPLC using solvent systems as previously described^{19,20}. BSF-1 concentration was determined by fluorescence analysis²¹ with bovine serum albumin as a standard. Natural and recombinant BSF-1 have identical specific activities ($2 \times 10^5 \text{ U } \mu\text{g}^{-1}$) as measured in the B-cell proliferation assay².

Table 2 Effects of BSF-1 and IL-2 on generation of cytolytic activity in allogeneic and syngeneic primary MLC

Culture supplement	Allogeneic MLC		Syngeneic MLC	
	Recovery (%)	LU per culture	Recovery (%)	LU per culture
Medium	55	4	21	<1
rIL-2 (ng ml^{-1})				
10^3	101	23	185	44
10^2	99	30	119	37
10	90	37	48	4
1	61	12	33	<2
10^{-1}	73	7	22	<1
10^{-2}	79	6	30	<1
rBSF-1 (ng ml^{-1})				
10^2	85	95	18	<1
10	65	124	15	<1
1	63	79	17	<1
10^{-1}	53	12	20	<1
10^{-2}	66	6	16	<1

LU, lytic units (see Fig. 1 legend). MLC were established with 2×10^6 C57BL/6 spleen cells and 5×10^6 irradiated (2500 R) spleen cells from either DBA/2 (left-hand data columns) or C57BL/6 (right-hand columns) and supplemented with the indicated dose of recombinant BSF-1 or IL-2. Lytic activity against ⁵¹Cr-labelled P815 was assessed on day 5, as in Fig. 1 legend. Spontaneous release was 125 c.p.m., maximum release, 886 c.p.m.

higher levels of CTL activity than control cultures, but the lytic activity was sevenfold less than that which developed in BSF-1 supplemented cultures. CTL generated in MLC supplemented with either BSF-1 or IL-2 were alloantigen-specific, because lytic activity directed against target cells syngeneic with the responding cell population was <2% of that directed against the specific allogeneic target (data not shown). These data indicate that BSF-1 is a potent helper factor for the generation of alloreactive CTL.

IL-2 is sufficient for the proliferation and differentiation of CTL from purified, resting precursors⁴. To determine whether BSF-1 can act on the CTL precursor-rich CD8⁺ (Lyt2⁺) lymphocyte subset in the absence of CD4⁺ (L3T4⁺) helper cells,

highly purified C57BL/6 CD8⁺ cells (>99% CD8⁺ as determined by flow cytometry) were cultured with T-cell depleted, irradiated, splenic stimulating cells in the presence of recombinant IL-2 or BSF-1. The CD8⁺ population gave rise to a high level of cytolytic activity in MLC supplemented with either IL-2 (192 lytic units (LU)—see Fig. 1 legend) or BSF-1 (714 LU) (Table 1). In addition, extensive cellular proliferation occurred in such cultures. These results demonstrate that CD8⁺ cells can, in the presence of exogenous BSF-1 and an appropriate antigenic stimulus, proliferate and acquire cytolytic activity independently of other T cells.

To test directly the relative efficiencies of BSF-1 and IL-2 in augmenting CTL generation, multiple concentrations of

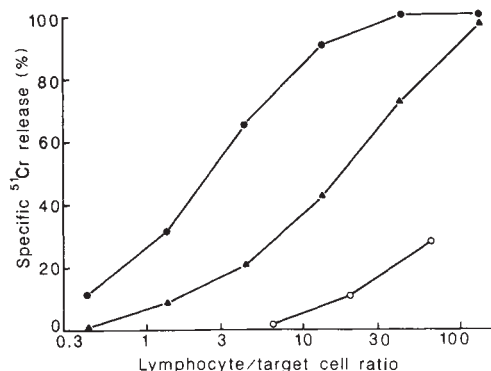


Fig. 1 Augmentation of CTL generation in primary MLC by BSF-1 and IL-2. Symbols are as follows: ○, medium only (92% recovery, <1.5 LU per culture); ●, culture supplemented with natural BSF-1 (186%, 176 LU per culture); ▲, culture supplemented with rIL-2 (186%, 23 LU per culture).

Methods. Mixed leukocyte cultures were established with 5×10^5 C57BL/6 spleen cells and 5×10^6 γ -irradiated (2,500 R from a ^{137}Cs source) DBA/2 splenic stimulating cells in Costar 16-mm diameter culture wells containing 2 ml Dulbecco's modified Eagle's medium (DMEM) containing 5% fetal bovine serum (FBS), 5×10^{-5} M 2-mercaptoethanol and additional amino acids²². Cultures were supplemented at initiation with homogeneous natural murine BSF-1 (nBSF-1) at 2 ng ml^{-1} purified as previously described², recombinant human IL-2 (rIL-2) at 10 ng ml^{-1} , or medium. Five days after culture initiation, lytic activity was tested by incubating, in duplicate 200 μl volumes, the indicated ratio of effector cells and ^{51}Cr -labelled P815 target cells (2×10^3 per well) in 96-well V-bottom microtitre plates. After a 3.5-h incubation, plates were centrifuged and 150 μl supernatant from each well were collected and counted in a γ -counter. Specific ^{51}Cr release is determined as $100 \times [\text{c.p.m. (experimental)} - \text{c.p.m. (spontaneous)}] / [\text{c.p.m. (maximum)} - \text{c.p.m. (spontaneous)}]$ where spontaneous release (118 c.p.m.) was determined by incubating P815 in medium, and maximum release (801 c.p.m.) by incubating P815 in 1N HCl. One lytic unit (LU) is defined as the number of cells required to achieve 50% lysis of 2×10^3 P815 target cells and is determined from the dose-response curve. Recovery was measured by cells recovered at day 5 as a percentage of the initial number of responding cells cultured.

homogeneous recombinant BSF-1 or IL-2 were added to primary MLC and the resultant lytic activity was measured five days later (Table 2). Although both lymphokines augmented cell proliferation and CTL activity, the levels of lytic activity that developed in cultures containing optimal doses of BSF-1 were approximately three- to fourfold higher than that observed in cultures supplemented with optimal doses of IL-2. In addition, at suboptimal lymphokine doses, approximately 10-fold less BSF-1 than IL-2 was required to obtain equivalent amounts of lytic activity (compare 12 LU per culture with 0.1 ng ml^{-1} rBSF-1 to 12 LU per culture with 1 ng ml^{-1} rIL-2). These data indicate that, in MLC established with this allogeneic strain combination, BSF-1 is an even more potent helper factor for the generation of CTL from unprimed precursors than is IL-2.

IL-2 can induce a lytic cell population *in vitro* in the absence of an overt antigenic stimulus⁵. To test whether BSF-1 has similar effects, cultures were established with C57BL/6 spleen cells and irradiated, syngeneic, C57BL/6 'stimulating' cells, supplemented with recombinant lymphokine and tested for cytotoxicity against P815 after five days. The results (Table 2) confirm previous reports that IL-2 can nonspecifically induce a lytic population; however, relatively high doses of IL-2 were required. In contrast to IL-2, BSF-1, even at equivalent high doses, induced no cell proliferation or anti-P815 killers in unstimulated cultures (right-hand columns in Table 2) in spite of its greater potency in augmenting the generation of CTL in allogeneic cultures (left-hand columns in Table 2) in the same experiment.

MLC populations cultured for long periods gradually lose

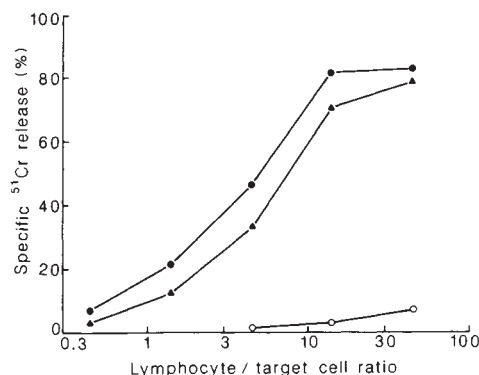


Fig. 2 Induction of cytolytic activity in long-term MLC by BSF-1 and IL-2. Mixed leukocyte cultures were established with 25×10^6 C57BL/6 spleen cells and 25×10^6 irradiated (2500R) DBA/2 splenic stimulating cells in 25 cm^2 flasks, 20 ml total volume. Fourteen days after initiation, cells were collected from primary cultures and 5×10^5 cells were recultured in Costar 16-mm culture wells in 2-ml volumes containing recombinant murine BSF-1 at 1 ng ml^{-1} , recombinant human IL-2 at 0.5 ng ml^{-1} , or medium. Three days later, culture contents were tested for lytic activity against ^{51}Cr -labelled P825 target cells as in Fig. 1 legend. Spontaneous release was 204 c.p.m., maximum release, 1,829 c.p.m. Culture supplements, with percentage recovery and activities are: ○, no supplement (45%, <0.6 LU per culture); ●, rBSF-1 (147%, 74 LU per culture); ▲, rIL-2 (180%, 59 LU per culture).

CTL activity, but can be re-induced to express high-level CTL activity by exposure to either allogeneic cells⁶ or appropriate culture supernatant⁷. To test the effects of BSF-1 on such MLC memory populations, cells obtained from day 14 C57BL/6 anti-DBA/2 primary MLC were cultured in the presence of recombinant IL-2 or BSF-1 and resultant cytotoxic activity was measured three days later. Exposure of the cells to either lymphokine resulted in cellular proliferation and the induction of high-level cytotoxic activity (Fig. 2) which was specific for the priming alloantigens (data not shown). Lytic activity generated in cultures incubated with BSF-1 was approximately 100-fold higher than that observed in control cultures incubated in medium only (Fig. 2), and 80-fold higher than the activity of the day-14 population before exposure to exogenous lymphokine (data not shown). Such a dramatic increase in lytic activity over three days strongly indicates a differentiative function for BSF-1, and raises the possibility that BSF-1 may be similar or identical to differentiation factors proposed by others⁸⁻¹³.

The data presented here show that BSF-1 induces both proliferation and cytotoxic activity in primary and memory MLC populations and thus demonstrate a novel regulatory mechanism for the generation of CTL. BSF-1 could act directly on activated CTL precursors or indirectly through a distinct T-cell type. Our data demonstrate that BSF-1 induces proliferation and cytotoxic activity in MLC established with CD8^+ cells in the absence of CD4^+ helper cells; thus, if a T-cell type other than the CTL precursor is involved in CTL generation in such cultures it must also be of the CD8^+ subset. In addition, it has recently been noted (K.H.G., L. Park, P. J. Morrissey, V. Price, D. L. Urdal and M.B.W., manuscript in preparation) that BSF-1 can drive proliferation of cloned CTL, helper T cells¹⁴ and activated CD4^+ and CD8^+ populations¹⁵. Taken together, these findings indicate that BSF-1 mediates its effects on CTL generation, at least in part, by its direct action on cells of the CTL lineage.

CTL induction by BSF-1 could involve an intermediary lymphokine such as IL-2. However T-cell proliferation induced by BSF-1 is not blocked by antibodies against IL-2 or its receptor^{14,15}. Furthermore, we have been unable to detect induction of IL-2 messenger RNA by BSF-1 in BSF-1-responsive T-cell clones or populations. Whereas these observations do not rule

out an involvement of IL-2 at some stage of BSF-1-augmented CTL generation, they do indicate that BSF-1 need not induce IL-2 to exert its effects.

The precise relationship of the T cells responsive to BSF-1 to those responsive to IL-2 has yet to be determined. We have recently obtained evidence indicating clonal heterogeneity of T cells in response to BSF-1, and the same clones exhibit concordance in response to IL-2 (K.H.G. *et al*, manuscript in preparation). Further experiments should elucidate the relationship between IL-2 and BSF-1 in terms of the cells they stimulate and the receptors through which their activities are mediated.

We thank Drs David Urdal and Virginia Price for the recombinant lymphokines, Robert Voice and Kurt Shanebeck for technical assistance, Drs Steven Gillis, Paul Conlon and Linda Park for reviewing the manuscript, and Linda Troup for preparation of the manuscript. M.B.W. is a Scholar of the Leukemia Society of America.

Received 22 January; accepted 3 March 1987.

- Howard, M. *et al*. *J. exp. Med.* **155**, 914-923 (1982).
- Grabstein, K. *et al*. *J. exp. Med.* **163**, 1405-1414 (1986).
- Lee, N. *et al*. *Proc. natn. Acad. Sci. U.S.A.* **83**, 2061-2065 (1986).
- Erard, J. *et al*. *J. Immun.* **134**, 1644-1652 (1985).
- Grimm, E. A. *et al*. *J. exp. Med.* **158**, 1356-1361 (1983).
- MacDonald, H. R., Engers, H. D., Cerottini, J.-C. & Brunner, K. T. *J. exp. Med.* **140**, 718-730 (1974).
- Ryser, J.-E., Cerottini, J.-C. & Brunner, K. T. *J. Immun.* **120**, 370-377 (1978).
- Raulet, D. H. & Bevan, M. J. *Nature* **296**, 754-757 (1982).
- Wagner, H. *et al*. *J. exp. Med.* **155**, 1876-1881 (1982).
- Garman, R. D. & Fan, D. P. *M. Immun.* **130**, 756-762 (1983).
- Finke, J. H., Scott, J., Gillis, S. & Hilfiker, M. L. *J. Immun.* **130**, 763-767 (1983).
- Kanagawa, O. *J. Immun.* **131**, 606-610 (1983).
- Mannel, D. N., Falk, W. & Droge, W. *J. Immun.* **130**, 2508-2510 (1983).
- Fernandez-Botran, R., Krammer, P. H., Diamantstein, T., Uhr, J. W. & Vitetta, E. S. *J. exp. Med.* **164**, 580-593 (1986).
- Hu-Li, J. *et al*. *J. exp. Med.* **165**, 157-172 (1987).
- Julius, M. H., Simpson, E. & Herzenberg, L. A. *Eur. J. Immun.* **3**, 645-649 (1973).
- Dialynas, D. P. *et al*. *Immun. Rev.* **74**, 29-56 (1983).
- Dennert, G., Hyman, R., Lesley, J. & Trowbridge, I. S. *Cell. Immun.* **53**, 350-364 (1980).
- Urdal, D. L. *et al*. *J. Chromatogr.* **296**, 171-179 (1984).
- Park, L. S., Friend, D., Grabstein, K. & Urdal, D. L. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
- Udenfriend, S. *et al*. *Science* **178**, 871-872 (1972).
- Cerottini, J.-C., Engers, H. D., MacDonald, H. R. & Brunner, K. T. *J. exp. Med.* **140**, 703-717 (1974).

CD4⁺ murine T cells develop from CD8⁺ precursors *in vivo*

Laurie Smith

ICRF Tumour Immunology Unit, Department of Zoology, University College London, Gower Street, London WC1E 6BT, UK

The adult murine thymus contains four subpopulations of thymocytes defined by the T-cell surface antigens CD4 (L3T4) (a marker of helper T cells) and CD8 (Lyt2) (a marker of cytotoxic/suppressor T cells): CD4⁺8⁻ and CD4⁺8⁺ (single positives), CD4⁺8⁺ (double positives) and CD4⁻8⁻ (double negatives). To understand how T cells develop in the thymus, it is important to determine the lineage relationships among these subpopulations. In particular, the status of double positives, which make up ~80% of the total thymocyte population¹, has long been controversial. Some propose that double positives are 'dead-end cells' that all die in the thymus, perhaps because they have been rejected by some selection process. Others suggest that, although most double positives die in the thymus, some develop into the more mature single positives that leave the thymus¹⁻⁵. The experiments presented here show that repeated injections of anti-CD8 monoclonal antibodies block the development of CD4⁺ cells, demonstrating that these cells develop from CD8⁺ precursors, probably double positive thymocytes, *in vivo*.

The experimental design is summarized in Fig. 1. Radiation bone marrow chimaeras were constructed in which a mixture

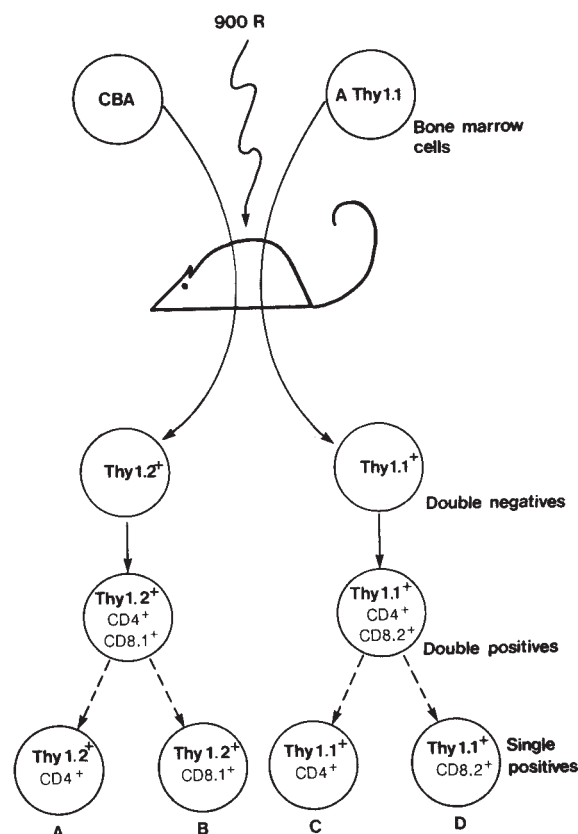


Fig. 1 Experimental design. In experiment 1, 900R-irradiated (⁶⁰Co source) (CBA×BALB/c) F₁ mice were injected intravenously 4-8 h post-irradiation with 5-10×10⁶ bone marrow cells, a 1:1 mixture of CBA (Thy-1.2, CD8.1) and AThy-1.1 (Thy-1.1, CD8.2) that had been T-cell depleted by treatment with the appropriate anti-Thy-1 antibodies and complement. Cell types expected to develop in the thymuses of these mice are shown. Solid arrows, known lineage relationships; dashed arrows, hypothetical lineage relationships. Mice were injected every 48 h intraperitoneally with either saline, anti-CD8.2 (clone 19/178 (ref. 14) mouse IgG_{2a}, provided by G. Hämmerling), or anti-CD8.1 (clone 49-11.1 (ref. 15) mouse IgG_{2a}, obtained originally from I. McKenzie). Antibodies were prepared by precipitation with 40% saturated ammonium sulphate from ascites fluid, and administered at an estimated 0.5-1.0 mg per mouse per injection. In experiment 2, hosts were (CBA×DBA/2)F₁ mice and bone marrow cell donors were AKR (Thy-1.1, CD8.1) and BALB/c (Thy-1.2, CD8.2).

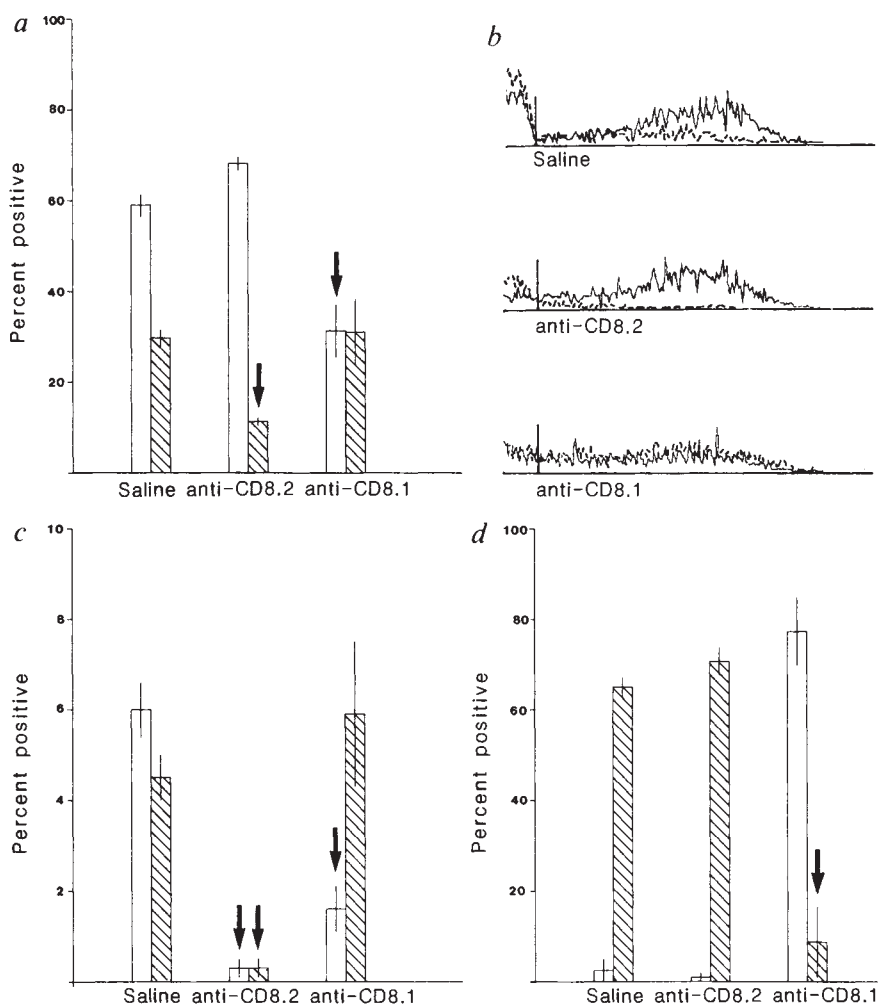
of cells bearing the two allelic forms of CD8, CD8.1 and CD8.2, would develop in parallel; these animals were injected continuously for 4-6 weeks following irradiation and bone marrow transfer with either anti-CD8.1 or anti-CD8.2 monoclonal antibodies (control chimaeras were injected with saline only). The bone marrow cell donors chosen also carried different alleles of Thy-1, providing the second marker necessary to trace the origin of CD4⁺8⁻ cells in these mixed chimaeras. Cells expected to develop in the thymuses of these mice are drawn as two separate streams, resulting in the production of four distinct sets of single positive cells (labelled A, B, C and D), which should be present in both the thymus and spleen. If CD4⁺8⁻ cells arise from CD4⁺8⁺ thymocytes, then depletion of Thy-1.2⁺ double positives with anti-CD8.1 would also be expected to deplete sets A and B but not C and D, whereas depletion of Thy-1.1⁺ double positives with anti-CD8.2 would also deplete sets C and D but not A and B. If CD4⁺8⁻ cells do not develop from CD8⁺ precursors, sets A and C should be unaffected by the antibody treatment. The experiment does not provide any information about the precursors of CD4⁺8⁺ cells, as these would be eliminated directly by anti-CD8 antibodies.

Fig. 2 Perturbation of CD4⁺ and CD8⁺ T cells by anti-CD8 antibody treatment. *a, d*, Mean proportion of CD4⁺ spleen T cells that are Thy-1.2⁺ (open bars) or Thy-1.1⁺ (hatched bars) from chimaeras injected with saline (left), anti-CD8.2 (centre) or anti-CD8.1 (right) in experiment 1, *d*, experiment 2. *b*, Histograms showing the relative proportions of Thy-1.2⁺ (solid line) and Thy-1.1⁺ (dashed line) CD4⁺ T cells of representative mice injected with saline (top), anti-CD8.2 (middle), or anti-CD8.1 (bottom). CD4 was labelled with fluorescein, and gates were set to include only CD4⁺ cells in the analysis of Thy-1.2 or Thy-1.1 labelled with phycoerythrin (same experiment as in *a*). *c*, Mean proportion of total spleen T cells that are Thy-1.2⁺CD8⁺ (open bars) and Thy-1.1⁺CD8⁺ (hatched bars) from chimaeras injected with saline, anti-CD8.2, or anti-CD8.1 (same experiment as *a*). Animals were killed 4 weeks (experiment 1) or 5 weeks (experiment 2) after irradiation and bone marrow transfer. In *a, c* and *d*, error bars indicate standard errors of means (three to six mice per group) and arrows indicate significant depletion of a cell population, relative to saline-injected controls, as determined by a Student's *t*-test ($p < 0.02$). CD4⁺ and CD8⁺ spleen T cells were always found to be less than 100% Thy-1⁺ both in these experiments and in normal mice. As shown in *c*, the intensity of Thy-1 staining of spleen T cells is very heterogeneous, and the amount of Thy-1 on a minority of T cells was presumably too low to be detected in these experiments.

Methods. Red blood cells were removed from spleen-cell suspensions from individual mice by controlled lysis. The cells were floated on Lympholyte-M (Cedarlane, relative density 1.087), yielding viable lymphocytes, and passed over nylon wool columns as described¹⁶ to enrich for T cells. Cells were incubated at 5×10^6 ml⁻¹ with saturating concentrations of monoclonal antibodies for 60 min at 4 °C, followed by a 30-min incubation at 4 °C with appropriately diluted second-step reagents. Reagents used for immunofluorescence were 3.168.8.1 (ref. 17; monoclonal rat IgM anti-CD8, non allele-specific), GK1.5 (ref. 18; monoclonal rat IgG anti-CD4), both provided by F. W. Fitch, biotin-conjugated H013.4 and H022.1 (ref. 19; monoclonal mouse IgM anti-Thy-1.2 and anti-Thy-1.1, respectively, obtained from ATCC; biotin conjugates prepared in this laboratory), fluorescein isothiocyanate (FITC)-conjugated MAR-18 (ref. 20; monoclonal anti-rat κ , Becton Dickinson) and phycoerythrin (PE)-conjugated streptavidin (Becton Dickinson). Dual-colour flow cytometry was performed on a fluorescence-activated cell sorter built at the Imperial Cancer Research Fund, gated to exclude non-viable cells. Control samples were incubated with fluorescent conjugates only. Further control staining confirmed that FITC-MAR-18 failed to bind to mouse anti-Thy-1 monoclonal antibodies and PE-streptavidin failed to bind to rat anti-CD4 and anti-CD8 monoclonal antibodies.

T cells from spleens of individual control and antibody-treated chimaeras were stained by indirect immunofluorescence with anti-CD4 or anti-CD8 in green and anti-Thy-1.2 or anti-Thy-1.1 in red, and the relative numbers of cells in sets A, B, C and D were determined by dual-colour flow cytometry. In Fig. 2*a*, the mean proportion of CD4⁺ T cells that are Thy-1.2⁺ or Thy-1.1⁺ is plotted for each group of mice in experiment 1. Saline-injected control chimaeras had both Thy-1.2⁺ CD4⁺ and Thy-1.1⁺ CD4⁺ T cells in their spleens. The arrows indicate that, as predicted from the hypothesis that CD4⁺CD8⁺ cells arise from CD4⁺CD8⁺ thymocytes, Thy-1.1⁺CD4⁺ T cells were significantly depleted in chimaeras injected with anti-CD8.2 and Thy-1.2⁺CD4⁺ T cells were significantly depleted in chimaeras injected with anti-CD8.1. In Fig. 2*b*, histograms showing Thy-1.2 and Thy-1.1 staining of CD4⁺ T cells from representative mice illustrate the results of experiment 1 directly.

In Fig. 2*c*, results for CD8⁺ T cells in experiment 1 are presented. Relative to control chimaeras, Thy-1.1⁺CD8⁺ (CD8.2⁺) cells are effectively depleted by anti-CD8.2 and Thy-1.2⁺CD8⁺ (CD8.1⁺) by anti-CD8.1. The unexpected finding that Thy-1.2⁺CD8⁺ (CD8.1⁺) T cells were also absent from anti-CD8.2 treated mice was reproducible (results for CD8⁺ T cells are



expressed as a proportion of the total T-cell population to illustrate this point). One possible explanation for this is based on the fact that the anti-CD8.2 antibody used is slightly cross-reactive on CBA cells, as shown by immunofluorescence and radioimmunoassays. This cross-reactive binding may be sufficient to eliminate CD8.1⁺ cells from peripheral lymphoid tissues, but not sufficient to perturb the development of CD8.1⁺ precursors of CD4⁺ cells in the thymus, an effect that depends on very high concentrations of antibody (data not shown).

Experiment 2 (Fig. 2*d*) shows that comparable results were obtained when a different pair of mouse strains were used as bone marrow cell donors. In this experiment, bone marrow donors were AKR (Thy-1.1, CD8.1) and BALB/c (Thy-1.2, CD8.2); thus the pairing of CD8 and Thy-1 alleles is reversed compared to experiment 1 (Fig. 1). Saline-injected controls were completely dominated by AKR-derived, Thy-1.1⁺ cells, which must have been more efficient at repopulating irradiated hosts in this experiment than BALB/c cells. Mice injected with anti-CD8.2 (BALB/c allele) are not significantly different from controls, as expected in this case because Thy-1.2⁺ cells were absent from control mice. But in chimaeras injected with anti-CD8.1 (AKR allele), Thy-1.1⁺CD4⁺ cells were dramatically depleted,

and replaced by Thy-1.2⁺CD4⁺ cells. CD8⁺ T cells were not directly examined in this experiment.

This demonstration that treatment of radiation bone marrow chimaeras with anti-CD8 monoclonal antibodies perturbs the development of CD4⁺ as well as CD8⁺ T cells provides direct evidence that CD4⁺ T cells develop from CD8⁺ precursors *in vivo*. The possibility that anti-CD8 antibody eliminates CD4⁺ cells directly by binding to a cross-reactive determinant present on these cells is extremely unlikely, as no such binding is detectable (data not shown), and the determinant would have to be polymorphic and tightly linked genetically to CD8. Other explanations for these results, such as that the development of CD4⁺ T cells depends somehow on the presence of CD8⁺ cells, or that antibody non-specifically interferes with the development of CD4⁺ T cells, can also be ruled out because of the experimental design: none of these alternatives explains how CD4⁺ cells of one strain but not the other could be affected by anti-CD8 antibody.

Certain considerations strongly suggest that the CD8⁺ precursors are thymic, rather than pre-thymic cells. During embryogenesis, the first haematopoietic cells to colonize the thymic epithelial rudiment express neither CD4 nor CD8^{6,7}. Such double negative cells, in both embryonic and adult thymus, have been shown to give rise intrathymically to CD4⁺8⁺, CD4⁺8⁻ and CD4⁻8⁺ cells⁸⁻¹⁰. Thus it seems unlikely that there would be a pre-thymic CD8⁺ precursor, and indeed, previous work has demonstrated that thymus-homing bone marrow cells are CD8⁻ (ref. 11). Therefore it seems most likely that the CD8⁺ precursors are double positive thymocytes, but the experiments presented here do not rule out the possibility that these precursors are CD4⁻. This possibility is raised by the observation that small numbers of functionally immature CD4⁺8⁺ (but not CD4⁺8⁻) thymocytes appear before CD4⁺8⁺ thymocytes during embryogenesis^{7,8}. Analysis of thymocytes from antibody-treated chimaeras is in progress, and should provide further evidence for the identity of CD8⁺ precursors of CD4⁺ T cells.

In the fifteen years since the original proposal that cortical thymocytes (double positives) mature in the thymus and migrate into the medulla (the location of most single positives)¹², many attempts have been made to provide direct evidence in support of this hypothesis. It has not been possible to demonstrate definitively that double positives are capable of giving rise to single positives either *in vitro*^{5,13} or *in vivo*¹⁰. The evidence presented here shows that a CD8⁺→CD4⁺8⁻ pathway does operate during T-cell ontogeny *in vivo*, but because elimination of CD4⁺8⁻ T cells by anti-CD8 antibody was never 100% efficient, it is not possible to conclude that this is the only pathway giving rise to this subclass of T cells. It would be interesting to examine similarly the precursors of CD4⁺8⁺ T cells using anti-CD4 antibody, but as yet, no allele-specific anti-CD4 monoclonal antibodies are available.

I thank N. A. Mitchison, M. J. H. Ratcliffe and D. B. G. Oliveira for helpful ideas and discussion, and M. C. Raff and N. A. Mitchison for reading the manuscript. Supported by a Marshall Scholarship from the Association of Commonwealth Universities.

Note added in proof: Evidence that the CD8⁺ precursor of CD4⁺ T cells is a thymic cell was obtained in a recent experiment: three weeks after irradiation and bone marrow transfer, anti-CD8 antibody had depleted double and single positive thymocytes, but not double negatives.

7. Ceredig, R., Dialynas, D. P., Fitch, F. W. & MacDonald, H. R. *J. exp. Med.* **158**, 1654-1671 (1983).
8. Kisielow, P., Leiserson, W. & von Boehmer, H. *J. Immun.* **133**, 1117-1123 (1984).
9. Kingston, R., Jenkinson, E. J. & Owen, J. J. T. *Nature* **317**, 811-813 (1985).
10. Fowlkes, B. J., Edison, L., Mathieson, B. J. & Chused, T. M. *J. exp. Med.* **162**, 802-822 (1985).
11. Lepault, F., Coffman, R. L. & Weissman, I. L. *J. Immun.* **131**, 64-69 (1983).
12. Weissman, I. L. *J. exp. Med.* **137**, 504-510 (1973).
13. Fowlkes, B. J., Edison, L., Mathieson, B. & Chused, T. M. in *Regulation of the Immune System* (eds Cantor, H., Chess, L. & Sercarz, E.) 275-293 (Liss, New York, 1984).
14. Hämmerling, G. J., Hämmerling, U. & Flaherty, L. *J. exp. Med.* **150**, 108-116 (1979).
15. Hogarth, P. M., Edwards, J., McKenzie, I. F. C., Goding, J. W. & Liew, F. W. *Immunology* **46**, 135-144 (1982).
16. Julius, M. H., Simpson, E. & Herzenberg, L. A. *Eur. J. Immun.* **3**, 645-649 (1973).
17. Sarmiento, M., Glasebrook, A. L. & Fitch, F. W. *J. Immun.* **125**, 2665-2672 (1980).
18. Dialynas, D. P. *et al. J. Immun.* **131**, 2445-2451 (1983).
19. Marshak-Rothstein, A. *et al. J. Immun.* **122**, 2491-2497 (1979).
20. Lanier, L. L., Gutman, G. A., Lewis, D. E., Griswold, S. T. & Warner, N. L. *Hybridoma* **1**, 125-129 (1982).

Two growth factor signalling pathways in fibroblasts distinguished by pertussis toxin

Jean Claude Chambard, Sonia Paris, Gilles L'Allemain & Jacques Pouyssegur

Centre de Biochimie, CNRS, Université de Nice, Parc Valrose, 06034 Nice, France

The primary action of a family of mitogens including bombesin^{1,2}, bradykinin³, vasopressin^{3,4} and α -thrombin⁵⁻⁷ is to activate the hydrolysis of polyphosphoinositides⁸. Hydrolysis of phosphatidylinositol 4,5-bisphosphate (PtdIns(4,5)P₂) by phospholipase C is mediated through coupling of surface receptors to a GTP-binding protein^{9,10} (G_p protein) which, in some cells^{6,11,12}, is inactivated by the toxin of *Bordetella pertussis*¹³. It is not known whether this signalling pathway is involved in initiating DNA replication, whereas it has been firmly established that reinitiation of DNA synthesis can be triggered without activation of PtdIns(4,5)P₂ hydrolysis by, for example, EGF (epidermal growth factor)^{14,15}, FGF (fibroblast growth factor)¹⁶ and insulin/IGF-I (insulin-like growth factor-I)¹⁴, members of a class of mitogens known to activate receptor tyrosine kinases^{17,18}. Taking advantage of the fact that Chinese hamster lung fibroblasts respond to either class of mitogens and that their G_p protein appears to be sensitive to pertussis toxin⁶, we have now analysed the toxin's effect on reinitiation of DNA synthesis and find that it inhibits up to 95% of thrombin-induced mitogenicity without affecting EGF- or FGF-induced DNA synthesis and proliferation. These findings strongly suggest that activation of PtdIns(4,5)P₂-phospholipase C has a determinant function in growth control, and confirm the existence of alternative growth factor-signalling pathways independent of polyphosphoinositide breakdown.

Figure 1 shows the effects of pertussis toxin on reinitiation of DNA synthesis in response to pure growth factors. Fewer than 1% of nuclei are labelled when growth-arrested CCL39 cells are incubated for 24 h with tritiated thymidine in serum-free medium (data not shown). Addition of α -thrombin alone at 1 nM or 10 nM induced DNA replication in 27% and 39% of the cells, respectively (Fig. 1a and b). In the presence of pertussis toxin, reinitiation of DNA synthesis was drastically inhibited because for the same thrombin concentrations the percentage of labelled nuclei dropped to 1.5% and 6% respectively (Fig. 1a', b'). In sharp contrast, reinitiation of DNA synthesis with a combination of mouse EGF¹⁹ and the basic form of bovine brain FGF²⁰ was not affected by pertussis toxin (Fig. 1c, c'). In more than 1,500 cells counted in three independent fields 31% and 32% of nuclei were labelled in the presence or absence of pertussis toxin, respectively. This result clearly demonstrates a differential inhibitory effect of pertussis toxin which is also

Received 18 December 1986; accepted 13 February 1987.

1. Scollay, R., Bartlett, P. & Shortman, K. *Immun. Rev.* **82**, 79-103 (1984).
2. Owen, J. J. T. in *Leukemia* (ed. Weissman, I. L.) 111-130 (Springer, Berlin, 1985).
3. Mathieson, B. J. & Fowlkes, B. J. *Immun. Rev.* **82**, 141-173 (1984).
4. Scollay, R. & Shortman, K. in *Recognition and Regulation in Cell-Mediated Immunity* (eds Watson, J. D. & Marbrook, J.) 3-30 (Dekker, New York, 1984).
5. Shortman, K. & Scollay, R. in *Recognition and Regulation in Cell-Mediated Immunity* (eds Watson, J. D. & Marbrook, J.) 31-60 (Dekker, New York, 1984).
6. Van Ewijk, W., Jenkinson, E. J. & Owen, J. J. T. *Eur. J. Immun.* **12**, 262-271 (1982).

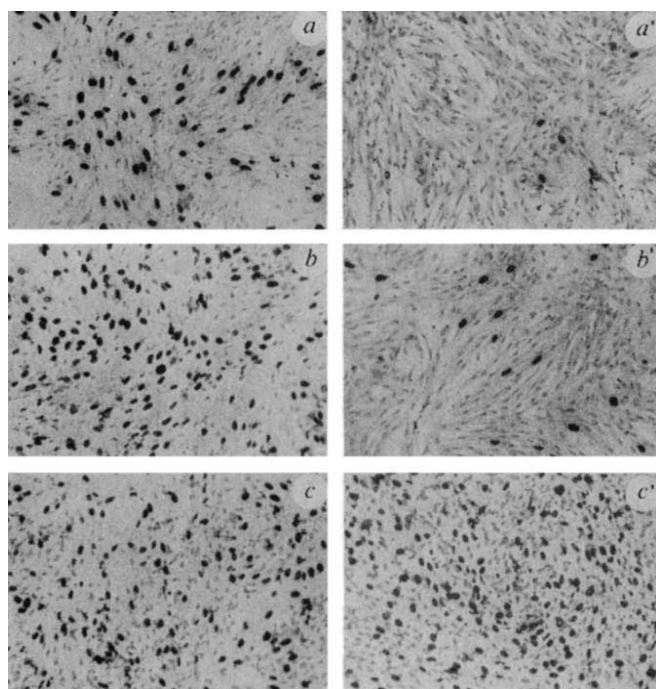


Fig. 1 Pertussis toxin specifically inhibits α -thrombin-induced reinitiation of DNA synthesis in CCL39 cells. Micrographs of autoradiograms showing labelled nuclei. Quiescent cultures of CCL39 cells in 35-mm-diameter dishes were reinitiated for 24 h by 1 nM α -thrombin (a, a') or 10 nM α -thrombin (b, b') or a mixture of 100 ng ml⁻¹ FGF and 50 ng ml⁻¹ EGF (c, c'). Pertussis toxin at 10 ng ml⁻¹ was absent (a, b, c) or present (a', b', c') during the complete period of reinitiation.

Methods. CCL39 cells cultivated in Dulbecco's modified Eagle's medium (DMEM) containing 5% fetal calf serum (FCS) were rendered quiescent by complete deprivation of serum for 24 h (ref. 33). Then cells were incubated for 27 h in a DMEM/Ham's medium (Gibco F12) mixture (1:1, v/v) containing 3 μ Ci ml⁻¹ methyl[³H]thymidine (1.5 μ M) and pertussis toxin at 10 ng ml⁻¹, where indicated. Growth factors were added for the last 24 h. Cell monolayers were rinsed with phosphate-buffered saline (PBS) and fixed with methanol. A thin layer of melted autoradiographic Ilford nuclear research emulsion K5 was then poured onto the dishes. After an exposure of 5 days, the autoradiograph was processed and cells were stained with Giemsa dye (Carlo Erba). Highly purified α -thrombin (2,660 NIH units mg⁻¹) was a gift from Dr J. W. Fenton II (New York State Department of Health). Mouse EGF and basic form of bovine brain FGF were purified to homogeneity in our laboratory by the methods of Savage and Cohen for EGF¹⁹ and Gospodarowicz *et al.* for basic FGF²⁰. Pertussis toxin (islet-activating protein) was purchased from List Biological Laboratories.

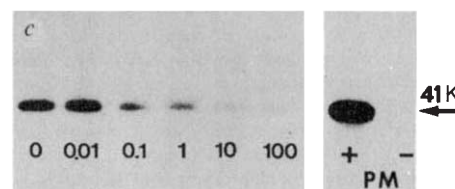
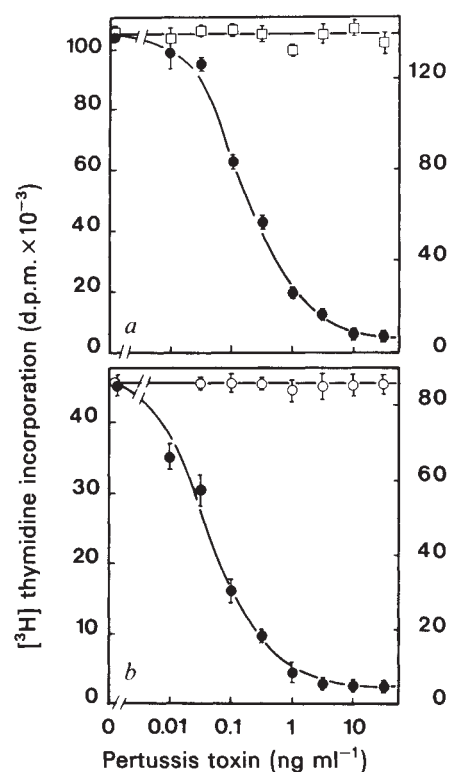


Fig. 2 ADP-ribosylation and inhibition by pertussis toxin of thrombin-induced reinitiation of DNA synthesis. Dose-response curves. a, b, Effect of pertussis toxin on growth-factor-induced DNA synthesis. G₀-arrested cultures of Chinese hamster lung fibroblasts (CCL39) (a) or secondary cultures of CHL fibroblasts (b) were preincubated for 4 h in a DMEM/Ham mixture (1:1, v/v) containing 1 μ Ci ml⁻¹ methyl[³H]thymidine (4.5 μ M) and the indicated concentrations of pertussis toxin. Incubation was continued for 24 h in the presence of either 1 nM α -thrombin (●) 100 ng ml⁻¹ basic FGF (□) or 50 ng ml⁻¹ EGF (○). Open symbols, right-hand vertical axis. c, Dose-response of toxin-catalysed *in vivo* ADP-ribosylation.

Methods. For a and b, CCL39 cells and CHL fibroblasts plated in 24-well multidishes (Nunc) were cultivated in DMEM supplemented with 5% FCS. Cells were rendered quiescent as in Fig. 1. At the end of the reinitiation period, cell monolayers were washed twice with PBS, fixed for 30 min with 10% ice-cold trichloroacetic acid and radioactivity associated with acid-insoluble material was measured by liquid scintillation spectrometry. Values represent average of duplicate determinations and correspond to the radioactivity incorporated per well. Maximal ³H-thymidine incorporation per well obtained with 10% FCS was 370 $\times 10^3$ d.p.m. and 145 $\times 10^3$ d.p.m. for CCL39 and CHL cells respectively.

For c and d, quiescent CCL39 cells were treated for 5 h with the indicated concentrations of pertussis toxin. Total cell homogenates were then incubated for 30 min at 37 °C with 5 μ g ml⁻¹ of activated pertussis toxin and 10 μ M of NAD⁺ (20 μ Ci ml⁻¹) in the conditions described³⁴, with minor modifications. Homogenates were solubilized in 2% SDS, boiled for 1 min and proteins separated by electrophoresis in 15% SDS-polyacrylamide gels (left-hand six lanes). In a separate experiment plasma membranes (PM) isolated from exponentially growing CCL39 cells were incubated under the same conditions as for the homogenates in the absence (-) or presence (+) of the toxin. Size determination used the 'low molecular weight' markers (Bio-Rad).

observed when cells are reinitiated either with EGF or FGF alone (Fig. 2). Secondary cultures of Chinese hamster lung (CHL) fibroblasts are more appropriate for comparing the mitogenicity of thrombin and EGF because EGF alone elicits a much stronger response in them than in the established cell line CCL39¹⁴. Whereas FGF- or EGF-induced DNA synthesis was remarkably unaffected by pertussis toxin up to 100 ng ml⁻¹, α -thrombin-induced DNA synthesis was found to be extremely sensitive to low concentrations of the toxin. Figure 2 shows that reinitiation of DNA synthesis elicited by 1 nM α -thrombin was almost completely inhibited in both CCL39 and CHL fibroblasts. About 95% inhibition was obtained at 3–10 ng ml⁻¹ and 50% inhibition at 0.1 ng ml⁻¹ of toxin. CCL39-derived plasma membranes, incubated with pertussis toxin and [³²P]NAD⁺ led to ADP-ribosylation of a prominent substrate of relative molecular mass (*M*_r) 41,000 (41K) (Fig. 2c). This protein probably represents the α -subunit of G_i¹³. Interestingly, the dose response

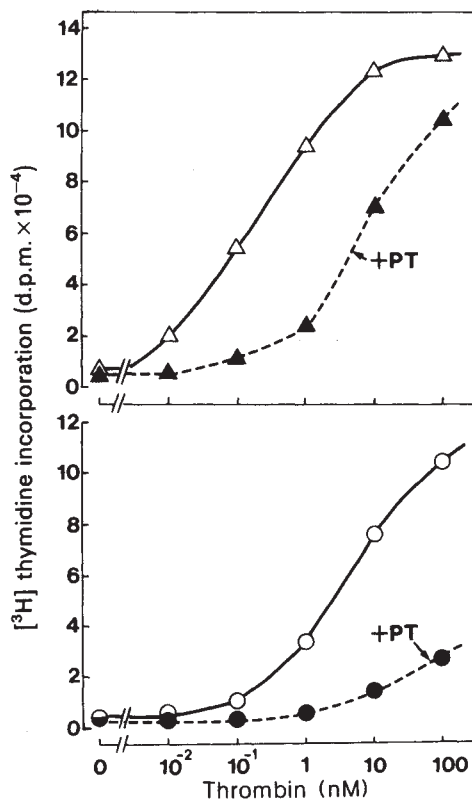


Fig. 3 The potentiating effect of insulin on thrombin-induced DNA synthesis is insensitive to pertussis toxin. G_0 -arrested CCL39 cells were reinitiated in absence (solid line) or presence (dashed line) of 10 ng ml^{-1} pertussis toxin as in Fig. 2. Insulin (10 $\mu\text{g ml}^{-1}$; Sigma) was added (upper panel only) together with the indicated concentrations of α -thrombin and incubated for 24 h at 37 °C. Values represent average of two separate experiments and correspond to the radioactivity incorporated per well.

of toxin-induced ADP-ribosylation of the 41K protein *in vivo*, parallels toxin-induced inhibition of DNA synthesis (Fig. 2c).

Another important class of growth-promoting agents for which it was of interest to analyse the effect of pertussis toxin are the co-mitogens insulin and IGF-I. The effects of insulin action cannot be tested directly because insulin alone fails to reinitiate DNA synthesis in G_0 -arrested CCL39 cells. Therefore we analysed whether pertussis toxin would prevent the potentiating action exerted by insulin when associated with other growth factors. Insulin, acting through IGF-I receptors in CCL39 cells²¹ strongly potentiates α -thrombin-induced DNA synthesis (compare upper and lower panels of Fig. 3). The potentiation is observed at all thrombin concentrations tested, with a more pronounced effect in the lower range of concentrations. Addition of pertussis toxin to thrombin-stimulated cells (lower panel of Fig. 3) shifted by 100-fold to the right the dose-response curve of α -thrombin. When insulin was added with thrombin (upper panel of Fig. 3), pertussis toxin still exerted a rightward shift of the dose-response curve of similar amplitude. This result clearly shows that the potentiating effect of insulin is not abolished by the action of pertussis toxin and that insulin does not modify the inhibitory effect of pertussis toxin on thrombin action.

The above inhibitory effects of pertussis toxin were restricted to α -thrombin action and to the capacity of the toxin to antagonize the transition from G_0 to S-phase. We also explored whether pertussis toxin could prevent cell proliferation in fetal calf serum or in growth factor-supplemented serum-free medium. Figure 4a shows that either basic FGF or α -thrombin as sole mitogen can sustain continuous cell proliferation, albeit at a low rate. Pertussis toxin did not affect the growth rate of FGF-stimulated cells but completely abolished growth and even decreased cell

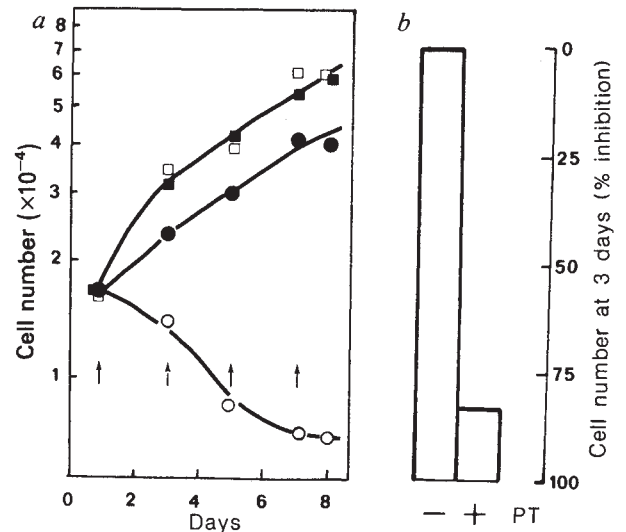


Fig. 4 Pertussis toxin inhibits thrombin- or serum-induced cell proliferation. **a**, CCL39 cells were plated at 1.5×10^4 cells per 35-mm dish in DMEM supplemented with 5% FCS. After 24 h, cultures were rinsed four times and incubated in serum-free DMEM containing 5 $\mu\text{g ml}^{-1}$ transferrin, 20 nM selenium supplemented with either 1 nM α -thrombin (○, ●) or 100 ng ml^{-1} basic FGF (□, ■). Pertussis toxin was added at 10 ng ml^{-1} to half the culture dishes (open symbols) and medium was changed every other day (arrows). Values represent average of duplicate cultures and correspond to total cell number per dish. **b**, CCL39 cells were plated at 5×10^3 cells per 35-mm dish in DMEM containing 5% FCS. One day later the medium was replaced with same medium with (+) or without (−) 10 ng ml^{-1} pertussis toxin. Results are presented as percentage inhibition of increase in cell number after 3 days.

viability in thrombin-supplemented medium. Finally, addition of pertussis toxin (10 ng ml^{-1}) to CCL39 cells growing exponentially in the presence of serum, reduced their growth rate by 83% (Fig. 4b), a finding previously reported for Swiss 3T3 cells²².

To sum up, pertussis toxin exerts a dose-dependent and drastic inhibitory effect on the reinitiation of DNA synthesis and cell growth promoted by α -thrombin and fetal calf serum, but does not affect the mitogenic action of EGF, FGF and insulin. How do we interpret this specific and differential effect?

A trivial explanation could be that pertussis toxin specifically interferes with the thrombin receptor of CCL39 cells. Indeed some odd effects have been directly attributed to protomer B of the toxin, for example mitogenicity in lymphocytes²³. We rule out this interpretation for the following reasons. (1) Pertussis toxin exerts similar inhibitory effects on serum-induced DNA synthesis. (2) Pertussis toxin also inhibits fluoroaluminate (AlF_4^-)-induced $\text{PtdIns}(4,5)\text{P}_2$ hydrolysis^{24,25}, an event not receptor-mediated, with the same concentration dependence. (3) Protomer B separated according to Tamura *et al.*²⁶ and used at a concentration up to 500 ng ml^{-1} has no effect on thrombin-induced DNA synthesis. (4) Toxin-induced ADP-ribosylation *in vivo* and inhibition of DNA synthesis follow the same dose responses (Fig. 2c).

From our studies of signal transduction in CHL fibroblasts, a more straightforward interpretation can be proposed. We have shown that α -thrombin rapidly induces the hydrolysis of $\text{PtdIns}(4,5)\text{P}_2$ in quiescent cells^{6,7}. Several lines of evidence strongly suggest that thrombin-induced activation of $\text{PtdIns}(4,5)\text{P}_2$ -phospholipase C is mediated through a new class of GTP-binding protein, G_p . (1) Pertussis toxin inhibits thrombin-induced $\text{PtdIns}(4,5)\text{P}_2$ hydrolysis and subsequent early mitogenic events^{6,25}. (2) AlF_4^- stimulates $\text{PtdIns}(4,5)\text{P}_2$ hydrolysis in intact cells in a pertussis toxin-dependent manner^{24,25}.

This phosphate analogue has been postulated to activate G proteins directly²⁷ by interacting with GDP at the nucleotide binding site²⁸. (3) GTP and nonhydrolysable GTP analogues directly stimulate basal and thrombin-induced PtdIns(4,5)P₂ hydrolysis in membranes of CCL39 cells²⁹. In sharp contrast to thrombin action, we found that EGF or EGF/insulin failed to stimulate hydrolysis of polyphosphoinositides¹⁴. Similarly, we demonstrated that basic or acidic forms of FGF do not mediate the activation of PtdIns(4,5)P₂-phospholipase C in CCL39 cells¹⁶. Therefore, the simplest interpretation of the differential inhibitory effect of pertussis toxin is that, by catalysing ADP-ribosylation of a G protein, it uncouples thrombin receptor from PtdIns(4,5)P₂-phospholipase C. It is not yet clear whether this inhibition results from direct modification of G_p or whether it is secondary to ADP-ribosylation of G_i. Nevertheless such a modification of G proteins is not expected to interfere with EGF-, FGF- and insulin-induced mitogenicity which is mediated through receptor tyrosine kinases^{17,18} and thus involves a separate route.

That α -thrombin and FGF activate in common very early mitogenic events^{16,25} (increase in cytoplasmic Ca²⁺ and pH; phosphorylation of the M_r 41/43K proteins³⁰, and S6; increased levels of c-fos and c-myc mRNA) suggests that the inhibitory effect of pertussis toxin is extremely specific. Our results lead us to conclude that hydrolysis of inositol lipids provides a signalling pathway of major importance to set in motion the G₀-to-S-phase transition and maintain the cells in the proliferative state, and that this pathway is not obligatory. Blocking it does not prevent mitogens acting through receptor tyrosine kinases to trigger DNA synthesis and cell division. A direct implication of these two proposals is that any agonist of inositol lipid breakdown should be considered as a potential mitogen. Indeed, some hormones, secretagogues and neurotransmitters are known to stimulate DNA synthesis. Another implication is that mutants deficient in PtdIns(4,5)P₂-phospholipase C or in receptor-phospholipase C coupling should not be lethal. Their isolation should provide valuable tools for studying the mechanism of phospholipase C activation. Finally, the use of pertussis toxin might be a useful approach to evaluating the function of the inositol lipid pathway in the process leading to growth factor relaxation and carcinogenesis in CHL fibroblasts^{31,32}.

We thank Miss M. Bonacci for secretarial assistance. This work was supported by grants from the Centre National de la Recherche Scientifique, the Institut National de la Santé et de la Recherche Médicale, the Laboratoires Glaxo, the Fondation pour la Recherche Médicale, and the Association pour la Recherche contre le Cancer.

Note added in proof: During the preparation of this paper a similar observation was reported for Swiss 3T3 in which pertussis toxin inhibits bombesin- but not PDGF-induced mitogenicity (Letterio, J., Coughlin, S. & Williams, L. *Science* **234**, 1117-1119 (1986)).

Received 28 October 1986; accepted 9 March 1987.

1. Brown, K., Blay, J., Irvine, R., Heslop, J. & Berridge, M. *Biochem. biophys. Res. Commun.* **123**, 377-384 (1984).
2. Zachary, I., Sinnen-Smith, J. & Rozengurt, E. *J. Cell Biol.* **102**, 2211-2222 (1986).
3. Vicentini, L. & Villereal, M. *Biochem. biophys. Res. Commun.* **123**, 663-670 (1984).
4. Kirk, C., Michell, R. & Hems, D. *Biochem. J.* **194**, 155-165 (1981).
5. Carney, D., Scott, D., Gordon, E. & La Belle, E. *Cell* **42**, 479-488 (1985).
6. Paris, S. & Pouyssegur, J. *EMBO J.* **5**, 55-60 (1986).
7. L'Allemain, G., Paris, S., Magnaldo, I. & Pouyssegur, J. *J. cell. Physiol.* **129**, 167-174 (1986).
8. Berridge, M. J. & Irvine, R. F. *Nature* **312**, 315-320 (1984).
9. Gomperts, B. *Nature* **306**, 64-66 (1983).
10. Haslam, R. & Davidson, M. *FEBS Lett.* **174**, 90-95 (1984).
11. Volpi, M. *et al. Proc. natn. Acad. Sci. U.S.A.* **82**, 2708-2712 (1985).
12. Ohta, H., Okajima, F. & Ui, M. *J. biol. Chem.* **260**, 15771-15780 (1985).
13. Ui, M. *Trends Pharmac. Sci.* **5**, 277-279 (1984).
14. L'Allemain, G. & Pouyssegur, J. *FEBS Lett.* **197**, 344-348 (1986).
15. Besterman, J., Watson, S. & Cuatrecasas, P. *J. biol. Chem.* **261**, 723-727 (1986).
16. Magnaldo, I. *et al. J. biol. Chem.* **261**, 16916-16922 (1986).
17. Hunter, T. & Cooper, J. A. *Rev. Biochem. Sci.* **54**, 897-930 (1985).
18. Huang, S. & Huang, J. *J. biol. Chem.* **261**, 9568-9571 (1986).
19. Savage, C. & Cohen, S. *J. biol. Chem.* **247**, 7609-7611 (1972).
20. Gospodarowicz, D., Cheng, J., Lui, G.-M., Baird, A. & Bohlen, P. *Proc. natn. Acad. Sci. U.S.A.* **81**, 6963-6967 (1984).

21. Van Obberghen-Schilling, E. & Pouyssegur, J. *Expl. Cell Res.* **147**, 369-378 (1983).
22. Hildebrandt, J., Stolzenberg, E. & Graves, J. *FEBS Lett.* **203**, 87-90 (1986).
23. Tamura, M. *et al. Biochemistry* **21**, 5516-5522 (1982).
24. Paris, S. & Pouyssegur, J. *J. biol. Chem.* **262**, 1970-1976 (1987).
25. Paris, S., Chambard, J. C., L'Allemain, G. & Pouyssegur, J. in *11th European Symposium on Hormones and Cell Regulation* (eds Dumont, J. E. & Nunez, J.) (in the press).
26. Tamura, M., Nogimori, K., Yajima, M., Ase, K. & Ui, M. *J. biol. Chem.* **258**, 6756-6761 (1983).
27. Gilman, A. *Cell* **36**, 577-579 (1984).
28. Bigay, J., Deterre, P., Pfister, C. & Chabre, M. *FEBS Lett.* **191**, 181-185 (1985).
29. Magnaldo, I., Talwar, H., Anderson, W. & Pouyssegur, J. *FEBS Lett.* **210**, 6-10 (1987).
30. Kohno, M. & Pouyssegur, J. *Biochem. J.* **238**, 451-457 (1986).
31. Pérez-Rodríguez, R., Chambard, J. C., Van Obberghen-Schilling, E., Franchi, A. & Pouyssegur, J. *J. cell. Physiol.* **109**, 387-396 (1981).
32. Renwick, D. E., Franchi, A., Pouyssegur, J. & Lagarde, A. E. *J. natn. Cancer Inst.* **77**, 105-114 (1986).
33. Pouyssegur, J., Franchi, A. & Silvestre, P. *Nature* **287**, 445-447 (1980).
34. Ribeiro-Neto, F. *et al. in Meth. Enzym.* **45**, 566-573 (1976).

Mammalian chiasma frequencies as a test of two theories of recombination

Austin Burt & Graham Bell

Department of Biology, McGill University, 1205 Avenue Docteur Penfield, Montreal, Quebec H3A 1B1, Canada

A broad survey of asexuality in the animal kingdom is sufficient to reject all theories of sex and recombination except two: the Red Queen and the Tangled Bank^{1,2}. The Red Queen theory states that an organism's biotic environment tends to be 'contrary'³, consistently evolving to the detriment of the organism; sex and recombination result in progeny genetically distinct from their parents and grandparents and thus less susceptible to the antagonistic advances made during the previous generations, particularly by their parasites^{4,5}. The alternative theory, the Tangled Bank, states that sex and recombination function to diversify the progeny from each other, thus reducing competition between them^{1,6}. An extensive survey of mammalian recombination shows that the total number of chiasmata in excess of one per bivalent is strongly correlated with generation time but uncorrelated with fecundity. We conclude that crossing-over may function to combat antagonists with short generation times but does not function to reduce sib competition. Chromosome number is selectively neutral with respect to these factors.

Our test derives from the opposing correlations between rates of recombination and life history parameters predicted by the two theories. According to the Red Queen theory, a coevolutionary model, generation time is an important determinant of selection for sex and recombination: the longer the time between generations, the more contrary will be the environment and the more intense the selection⁷⁻⁹. Thus it predicts that recombination is positively correlated with age to maturity (an index of generation time). However, the Tangled Bank is a model of sib competition and selection for sex and recombination will be influenced by fecundity: the larger the number of offspring, the higher the levels of recombination necessary for sufficient diversification¹⁰. Thus it predicts a positive correlation between recombination and number of offspring. Mammals provide a critical test for the theories because age to maturity correlates negatively with litter size¹¹ and thus at least one of the specific predictions will be rejected.

The standard measure of recombination is the recombination index $RI = n + TC$, where n is the haploid number of chromosomes and TC is the total number of chiasmata. In our data set of nondomesticated male mammals, $RI = 47 \pm 3.2$ (mean $\pm$ s.e.m., $N = 32$). However, there are a number of constraints preventing either of the two components of this value from being determined solely by selection for recombination. Changes in haploid number: (1) usually have (deleterious) phenotypic effects on the organism¹² which might obscure any effect on recombination; (2) often will not spread through a population simply due to negative heterosis^{13,14}; and (3) may act as isolating

Table 1 Correlation statistics for measures of recombination and life history in mammals

Variable (1)	Variable (2)	Partial correlate	Correlation statistic	N	P
EC	n	—	0.001	32	$P > 0.9$
$\log_{10}$ MAT	EC	—	0.875	24	$P < 0.001$
$\log_{10}$ MAT	EC	$\log_{10}$ WT	0.689	24	$P < 0.001$
$\log_{10}$ MAT	RI	—	0.471	24	$0.05 > P > 0.02$
$\log_{10}$ MAT	n	—	0.107	24	$0.9 > P > 0.5$
$\log_{10}$ LS	EC	—	-0.504*	24	$P < 0.01$
$\log_{10}$ LS	RI	—	-0.332*	24	$0.05 > P > 0.01$
$\log_{10}$ WT	EC	—	0.744	24	$P < 0.001$
$\log_{10}$ WT	$\log_{10}$ (EC/MAT)	—	-0.675	24	$P < 0.001$
$\log_{10}$ WT	EC	$\log_{10}$ MAT	0.021	24	$P > 0.9$

Variables as follows: EC, excess chiasma frequency; n, haploid chromosome number; MAT, age at maturity; WT, adult body weight; RI, recombination index; LS, litter size. Correlation statistics are Pearson correlation coefficients, except * which are Kendall's τ ; N, no. of species.

Table 2 Least-squares regression statistics for equations of the form $Y = a + bX$ ($N = 24$)

Y	X	a	b	$\bar{X}$	Σx^2	s^2_{YX}
EC	$\log_{10}$ MAT (d)	-14.4	10.8	2.5	9.969	15.953
EC	$\log_{10}$ WT (kg)	13.3	5.6	-0.21	26.096	30.458
$\log_{10}$ (EC/MAT)	$\log_{10}$ WT (kg)	-1.56	-0.30	-0.21	26.096	0.124

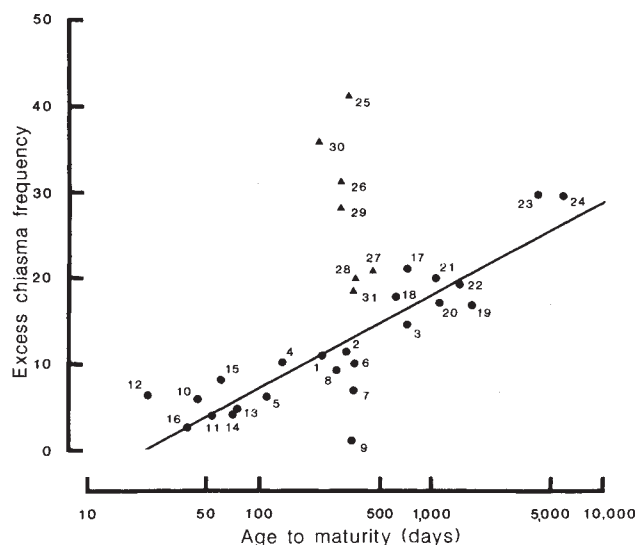


Fig. 1 Semilog plot of mammalian excess chiasma frequencies (EC) as a function of age to maturity. Only male mammals were considered. Circles, nondomesticated species; solid line, least squares regression. Triangles, domesticated species not included in the regression. 1, *Dasyuroides byrnei*; 2, *Dasyurus viverrinus*; 3, *Sarcophilus harrisii*; 4, *Sminthopsis crassicaudata*; 5, *Perameles gunnii*; 6, *Isodon macrourus*; 7, *Dasyurus novemcinctus*; 8, *Oryctolagus cuniculus*; 9, *Thomomys bottae*; 10, *Cricetus cricetus*; 11, *Mesocricetus auratus*; 12, *Lagurus lagurus*; 13, *Meriones unguiculatus*; 14, *Apodemus sylvaticus*; 15, *Rattus norvegicus*; 16, *Mus musculus*; 17, *Cebuella pygmaea*; 18, *Saguinus oedipus*; 19, *Macaca fuscata*; 20, *Macaca mulatta*; 21, *Macaca nemestrina*; 22, *Cynopithecus niger*; 23, *Pan troglodytes*; 24, *Homo sapiens*; 25, *Canis familiaris*; 26, *Felis catus*; 27, *Sus scrofa*; 28, *Bos taurus*; 29, *Capra hircus*; 30, *Ovis aries*; 31, *Equus caballus*.

mechanisms between populations^{13,14}, resulting in selection that operates between groups rather than individuals and consequently is much weaker¹⁵. Chiasma frequencies are constrained by the fact that crossing-over is often required for proper segregation of autosomes¹⁶. Meiotic autosomes lacking a chiasma (univalents) are very rare in mammals (for example, 0.2% for wild *Mus musculus*¹⁷, 0.05% for *Homo sapiens*¹⁸). Furthermore, for all 21 species reviewed in this study for which sufficient data were available, every autosomal bivalent had at

least one chiasma, though XY bivalents may have none. These constraints suggest an alternative measure of recombination, the excess chiasma frequency (EC), defined as the number of chiasmata in excess of one, summed across bivalents. This trait does not have the constraints associated with the recombination index, which now can be calculated as $RI = 2n + EC$ (assuming the sex bivalent has at least one chiasma). Excess chiasma frequency averages 11.6 ± 1.32 (mean $\pm$ s.e.m., $N = 32$) and is independent of haploid number (Table 1). Thus the component of the recombination index most likely to be responsive to selection for recombination is unaffected by the other major component. With this unexpected observation in mind, we now turn to the critical test of the theories, using both recombination index and excess chiasma frequency as measures of recombination.

Complete life history data were available for 24 species. Excess chiasma frequency correlates very well with age to maturity (Tables 1 and 2, Fig. 1). Partial correlation with body weight shows that this association is robust as well as tight: mammals that mature late for their size also have many excess chiasmata for their size (Table 1). Recombination index correlated less well with age to maturity and the association is wholly due to the above relation, as haploid number is independent of age to maturity. Finally, neither excess chiasma frequency nor recombination index correlates positively with litter size.

Because age to maturity correlates with body size¹⁹, the allometric equation relating excess chiasma frequency to body weight is also significant (Tables 1 and 2). This observation parallels others that sexuality, broadly defined, tends to increase with increasing body size^{1,6,20}. However, recombination per unit time decreases with increasing body size and size has no effect on excess chiasma frequency independently of age to maturity (Table 1). This latter result contradicts one author's prediction⁹; he had reasoned that the evolution of parasites should be faster on larger hosts than on smaller ones.

All the above analyses involve only species that have not been intensively bred; domesticated mammals tend to have much higher excess chiasma frequencies for their age to maturity (residual = 16 ± 3.6 , mean $\pm$ s.e.m., $N = 7$, Fig. 1). This observation parallels a similar finding in cultivated strains of the rye grass *Lolium perenne*²¹. We suggest that high rates of recombination are indirectly selected in breeding programmes because of their effect in removing negative correlations between desirable

characters²². Put another way, high recombination is an adaptation to an environment characterized by intense selection in small populations for novel combinations of traits²³, exemplifying lottery model of selection²⁴. Apart from its intrinsic interest, this result also demonstrates that excess chiasma frequencies can change dramatically in only ten thousand years. Thus, excess chiasma frequencies observed at present probably reflect very recent patterns of natural selection and phylogenetic constraints are likely to be minimal.

To conclude, the Red Queen theory has revealed a simple relation which accounts for 75% of the variance in excess chiasma frequency in mammals, although the effect of generation time can be swamped by the intense artificial selection of breeding programmes. The results are also suggestive of a division of labour between chiasmata, one per bivalent to meet the mechanical exigencies of proper segregation and the others for recombination. Finally, chromosome number does not seem to be under selection for its effects on recombination, nor does it affect selection for recombination.

Received 15 December 1986; accepted 9 February 1987.

1. Bell, G. *The Masterpiece of Nature: The Evolution and Genetics of Sexuality* (Croom Helm, London, 1982).
2. Bell, G. *Experientia* **41**, 1235-1245 (1985).
3. Trivers, R. L. *Social Evolution* (Cummings, Menlo Park, 1985).
4. Jaenike, J. *Evol. Theory* **3**, 191-194 (1978).
5. Hamilton, W. D. *Oikos* **35**, 282-290 (1980).
6. Ghiselin, M. T. *The Economy of Nature and the Evolution of Sex* (University of California, Berkeley, 1974).
7. Bremmnerman, H. J. *J. theor. Biol.* **87**, 671-702 (1980).
8. Hamilton, W. D. in *Population Biology of Infectious Disease Agents* (eds Anderson, R. M. & May, R. M.) 269-296 (Springer, Berlin, 1982).
9. Tooby, J. *J. theor. Biol.* **97**, 557-576 (1982).
10. Maynard Smith, J. *J. theor. Biol.* **63**, 245-258 (1976).
11. Eisenberg, J. F. *The Mammalian Radiations: An Analysis of Trends in Evolution, Adaptation, and Behaviour* (University of Chicago, 1981).
12. Patton, J. L. & Sherwood, S. W. *A. Rev. Ecol. Syst.* **14**, 139-158 (1983).
13. White, M. J. D. in *Cyrtotaxonomy and Vertebrate Evolution* (eds Chiarelli, A. B. & Capanna, E.) 95-128 (Academic, London, 1973).
14. Wilson, A. C., Bush, G. L., Case, S. M. & King, M.-C. *Proc. natn. Acad. Sci. U.S.A.* **72**, 5061-5065 (1975).
15. Williams, G. C. *Adaptation and Natural Selection: A Critique of Some Current Evolutionary Thought* (Princeton University, 1966).
16. Mather, K. *Biol. Rev.* **13**, 252-292 (1938).
17. Searle, A. G., Berry, R. J. & Beechey, C. V. *Mutat. Res.* **9**, 137-140 (1970).
18. Hultén, M. *Hereditas* **76**, 55-78 (1974).
19. Blueweiss, L. et al. *Oecologia* **37**, 257-272 (1978).
20. Bell, G. in *The Origin and Evolution of Sex* (eds Halvorson, H. O., Sager, R., Monroy, A. & Segal, S.) 221-256 (Liss, New York, 1985).
21. Rees, H. & Dale, P. J. *Chromosoma* **47**, 335-351 (1974).
22. Hill, W. G. & Robertson, A. *Genet. Res.* **8**, 269-294 (1966).
23. Thompson, V. *Evol. Theory*, **1**, 131-156 (1976).
24. Williams, G. C. *Sex and Evolution* (Princeton University, 1975).

Sliding movement of single actin filaments on one-headed myosin filaments

Yoshie Harada, Akira Noguchi, Akiyoshi Kishino & Toshio Yanagida

Department of Biophysical Engineering, Faculty of Engineering Science, Osaka University, Toyonaka, Osaka, Japan

The myosin molecule consists of two heads, each of which contains an enzymatic active site and an actin-binding site. The fundamental problem of whether the two heads function independently or cooperatively during muscle contraction has been studied by methods using an actomyosin thread¹, superprecipitation²⁻⁴ and chemical modification of muscle fibres⁵. No clear conclusion has yet been reached. We have approached this question using an assay system in which sliding movements of fluorescently labelled single actin filaments along myosin filaments can be observed directly^{6,7}. Here, we report direct measurement of the sliding of single actin filaments along one-headed myosin filaments in which the density of heads was varied over a wide range. Our results show that cooperative interaction between the two heads of myosin is not essential for inducing the sliding movement of actin filaments.

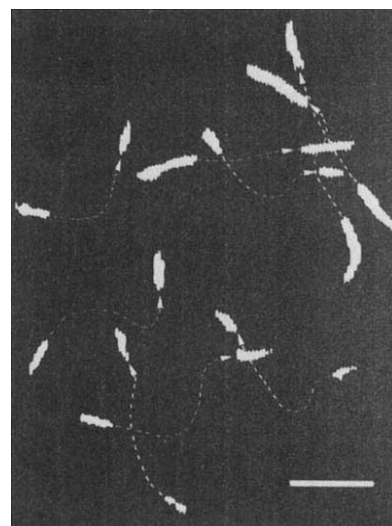


Fig. 1 Sliding movement of fluorescently labelled single actin filaments on a glass slide coated with single-headed myosin filaments. To demonstrate the movement of the actin filaments, two video images taken 1.5 s apart were photographed on the same frame of a film by double exposure. Dotted lines, traces of the movement of actin filaments; arrowheads, direction of movement. Scale bar, 5 μ m.

Methods. A glass slide was coated with myosin filaments by cleaning it in ethanol and then exposing it to buffer 1 (50 mM KCl, 10 mM HEPES pH 7.0, 1 mM MgCl₂, 1 mM dithiothreitol (DTT)) containing myosin filaments. Dark-field microscopy indicated that it is important for the fast movement of actin filaments that the surface of coverslip is covered by a continuous thin layer of myosin filaments. To prepare such a surface, the concentration of myosin filaments and the incubation time were appropriately changed from 2 to 0.5 mg ml⁻¹ and from 1 to 30 min at 0 °C respectively, monitoring the surface by dark-field microscopy¹⁴. The unbound filaments were washed away with buffer 2 (10 mM HEPES pH 7.0, 5 mM MgCl₂, 0.2 mM CaCl₂, 1% 2-mercaptoethanol and 2 mg ml⁻¹ bovine serum albumin (BSA)), then the slide was exposed to fluorescently labelled actin filaments (0.5 μ g ml⁻¹) suspended in test solution (50 mM KCl, 20 mM HEPES pH 7.8, 5 mM MgCl₂, 0.2 mM CaCl₂, 3 mM ATP and 1% 2-mercaptoethanol). Movements of actin filaments were observed at 23 °C. Actin was obtained from rabbit skeletal muscle and purified by the method of Spudich and Watt¹⁵. Actin filaments were labelled with phalloidin-tetramethyl-rhodamine by a method described previously⁸. Single fluorescent actin filaments were observed with an inverted microscope (Zeiss IM35) equipped with epifluorescence optics, a Zeiss Neofluor $\times 100$ object (oil immersion; NA 1.3), a 100-W mercury arc lamp and a Zeiss rhodamine filter (excitation, 554 nm; emission >580 nm). Images were recorded on videotape with a high sensitivity television camera (Ikegami SIT camera CTC-9000) and a video recorder (Victor CR 6650L)⁸.

We have demonstrated that single actin filaments labelled with a complex of fluorescent dye and phalloidin, which stabilizes the filament structure of actin, can be clearly and continuously seen under the fluorescence microscope⁸. Using this technique, Kron and Spudich have recently developed a new movement assay system⁶, which we have used to measure the sliding movement of actin filaments along myosin filaments directly. Fluorescently labelled actin filaments were added to a glass slide coated with myosin filaments and the movements of actin filaments were observed by fluorescence microscopy (Fig. 1). Most rapid movement of actin filaments along double-headed myosin filaments was observed in a buffer containing 50 mM KCl and 3 mM ATP. When the concentration of KCl was decreased from 50 to 0 mM, the average velocity of the filaments decreased from 7.2 to 5.1 μ m s⁻¹ at 23 °C, which agrees with the results obtained for chemically skinned muscle fibres⁹. At more than 75 mM KCl, the number of moving actin filaments was greatly decreased,

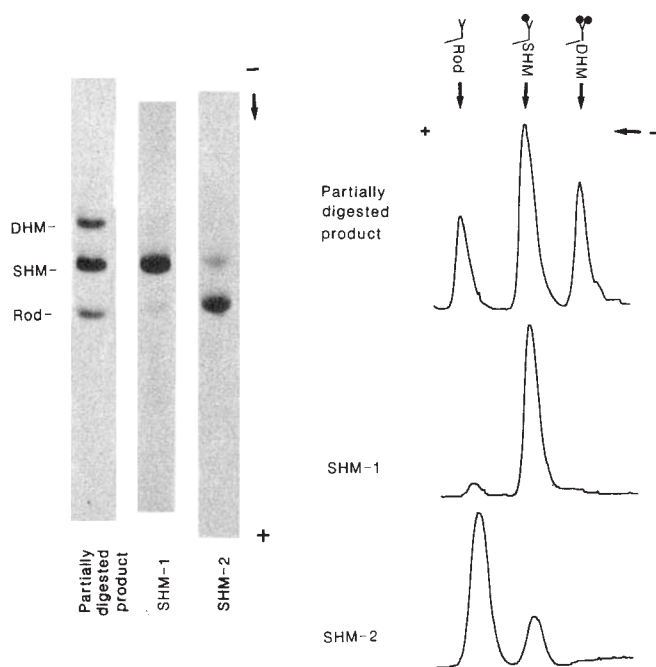


Fig. 2 Preparations of single-headed myosin filaments. In the first preparation (SHM-1), myosin filaments were digested by papain³ until double-headed myosin was barely detectable on Coomassie blue stains of polyacrylamide gels. Then single-headed myosin was purified by the methods of Cooke and Franks¹. The short thick filaments were formed by rapidly suspending the purified preparation of single-headed myosin, dissolved in a high ionic solution, in buffer 1¹⁶. Purity was checked by 3% PAGE in the presence of 20 mM inorganic pyrophosphate (PPi)¹. Gel patterns of typical samples are shown; the sample contained about 90% single-headed myosin and 8% rods in a molar ratio, and a small amount of double-headed myosin. The amount of double-headed myosin was estimated by comparing the pattern of the sample with those of the sample plus 0.5–2% of double-headed myosin and found to be <2%, probably 1%. It has been shown that concentrations of the myosin species determined by this gel method agree reasonably well with those from electron microscopy¹. Myosin was obtained from rabbit skeletal muscle and purified by a method described previously⁸. In the second preparation (SHM-2), >90% of the myosin heads were removed from synthetic myosin filaments, which had been formed by rapid dilution from 0.6 M KCl, by papain digestion. The digested heads were removed and the treated filaments used without a cycle of dissolution and subsequent filament formation. The gel of a typical sample of SHM-2 had 15% single-headed myosin and 85% rod (molar ratio). The line corresponding to double-headed myosin was barely visible on the gel even if the sample was overloaded; its content, therefore, was estimated at <0.5%.

and the sliding movement was no longer observed at KCl concentrations greater than 100 mM.

Single-headed myosin, obtained by papain digestion³ and purified by the method of Cooke and Franks¹, was capable of forming filaments (Fig. 2). The sliding velocity of actin filaments along thick filaments formed from the purified single-headed myosin (SHM-1) or double-headed myosin (DHM) is similar in the range 0–30 mM KCl (Fig. 3). At KCl concentrations over 30 mM, few actin filaments interact with the thick filaments in the presence of 3 mM ATP. When the concentration of ATP is decreased to 100 μ M, however, the sliding movement is observed even at 75 mM KCl. The preparations of single-headed myosin used in this study contained about 2% molar ratio of double-headed myosin (Fig. 2). We tested the possibility that the contaminating double-headed myosin is responsible for the movement of actin filaments by observing the movement of actin filaments on glass coated with myosin filaments that had been

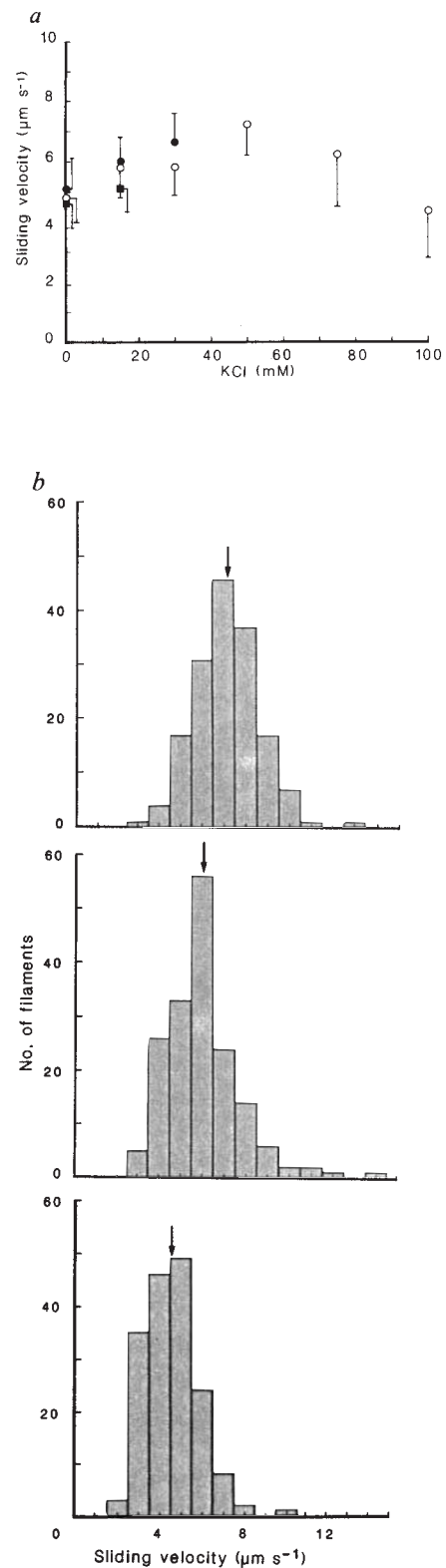


Fig. 3 Sliding velocities at various concentrations of KCl. *a*, The sliding velocities of actin filaments were measured in the test solution with the KCl concentration varied from 0 to 100 mM at 23 °C. The number of filaments that travelled more than 5 μ m were counted using double-headed myosin (○), single-headed myosin (SHM-1) (●) or single-headed myosin (SHM-2) (■). Error bars, standard deviations for samples of 50 different actin filaments. *b*, Histograms of sliding velocities of actin filaments running on double-headed myosin filaments at 50 mM KCl (above), SHM-1 at 15 mM KCl (middle) and SHM-2 at 0 mM KCl (below). Arrows, average velocities.

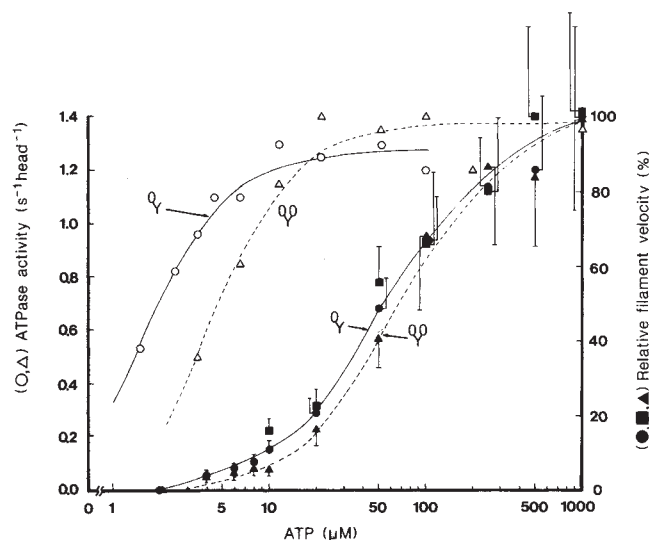


Fig. 4 Relative velocities of actin filaments (filled symbols) and ATPase activities of actomyosin (open symbols) plotted as a function of the ATP concentration. The velocities of actin filaments along double-headed myosin filaments ($\blacktriangle$), single-headed myosin ($\bullet$, SHM-1; $\blacksquare$, SHM-2) were measured in the test solution containing 15 mM KCl, 2 mM phosphocreatine and 0.5 mg ml⁻¹ creatine kinase (Sigma; 160 U mg⁻¹ at 30 °C) at various ATP concentration between 0 to 1 mM at 23 °C. Error bars, standard deviations for 30 different filaments. ATPase activity was measured under the same conditions and determined by the rate of creatine release¹⁷. The protein concentration of actin, double-headed myosin (Δ) and single-headed myosin ($\circ$) used for ATPase rate measurements were 0.5 mg ml⁻¹, 0.05 mg ml⁻¹ and 0.05 mg ml⁻¹ respectively.

prepared by three methods: (1) rods without heads (obtained by papain digestion of myosin) and double-headed myosin were mixed in a molar ratio of 98:2 and the hybrid filaments, formed by decreasing the ionic strength, were added to the coverslips; and (2) the mixture of rods and intact myosin in high-ionic-strength medium were added to coverslips and filaments formed on the surface; and (3) a mixture of filaments consisting of rod (98% molar ratio) and double-headed myosin (2%) which had been formed separately were added to the coverslip. In all three experiments, no movement of actin filaments was observed, indicating that the contaminating double-headed myosin was not responsible for the movement. Furthermore, the movement of actin was observed even though the content of contaminant double-headed myosin in a mixture of single-headed myosin and rods (SHM-2) was <0.5%. These results suggest that the double-headed structure of myosin is useful for increasing the affinity of myosin for actin but cooperative interaction between the two heads is not essential to the development of the sliding force.

But this result does not exclude the possibility that the original cooperative interaction between the two heads of myosin is replaced by interactions between two adjacent single heads, because the density of heads on the thick filaments was very high. To examine this possibility we measured the sliding movement of actin filaments along thick filaments which contained a low density of single heads. Low-density single-headed-filament surfaces were prepared by thinning >90% of the heads from myosin filaments by papain digestion. Figure 2 shows the pattern on polyacrylamide gel electrophoresis (PAGE) of the sample (SHM-2) and its densitometer scan. The gel pattern shows that the treated filaments contain 85% rods and 15% single-headed myosin (molar ratio). The band corresponding to double-headed myosin is very faint, corresponding to <0.5% of the total. The actin filaments were seen to move along SHM-2 filaments at almost the same velocity as that before thinning of the heads (in the range 0–15 mM KCl) (Fig. 3). If we assume

that the synthetic filament, like intact thick filaments, contains three myosin molecules per 14.3 nm, then the average interval between adjacent heads along the filament would be about 32 nm (3 myosin molecules $\times$ 0.15/14.3 nm = 1 single-headed myosin/32 nm). This almost completely excludes the possibility of pairing of adjacent single heads on the filament, assuming random papain digestion. Because short actin filaments no more than 1 μ m long move a long distance along these filaments at 5 μ m s⁻¹ or faster, the result suggests that a very few myosin heads are sufficient to achieve the maximum velocity. Detailed experiments to obtain the lower limit of the number of myosin heads required for this movement will provide important information about the length of the sliding distance of an actin filament during one ATP cycle—a distance that may be as large as 60 nm under zero-loaded conditions^{10,11}. When the content of single-headed myosin was decreased to less than 15%, both the velocity of actin filaments and the number of moving filaments are decreased; at <5%, moving filaments were no longer observed. This result agrees with previous studies in that a mixture of double-headed myosin and rods in a molar ratio of 2:98 cannot support the movement of actin filaments. These results support the conclusion that cooperative interaction between the two heads of myosin is not essential for inducing the sliding movement of actin filaments.

The double-headed structure of myosin, however, may be important for generating a large force with high efficiency. The above results do not address this question because the sliding movement was measured under zero-loaded conditions. Because it is difficult to provide load directly to the movement of actin filament in our assay, we examined the properties of relative movement between actin and myosin filaments at low concentrations of ATP. Figure 4 shows the dependence of the velocity of actin filaments and of actomyosin ATPase activity on the ATP concentration. When double-headed myosin was used, the half-maximal velocity occurs at $\sim$ 60 μ M ATP as previously reported⁶, whereas the half-maximal actomyosin ATPase activity is at $\sim$ 4 μ M. This large difference is expected because the rigid crossbridges lacking nucleotide that form at low ATP concentration provide resistance against the sliding force developed by active crossbridges, so that the velocity is determined by the balance between resistant and propulsive forces¹². For example, the fraction of rigid crossbridges lacking nucleotide has been estimated at $\sim$ 10% at 60 μ M ATP¹³, so that the net force between the 10% rigid crossbridges and 90% active crossbridges results in the half-maximal velocity. Single-headed myosin shows a similar dependence of the sliding velocity on the concentration of ATP, suggesting that it can develop sufficient sliding force to overcome the resistant forces of rigid crossbridges. This is consistent with previous results obtained using the actomyosin thread assay system¹. However, whereas the concentration of ATP that produces half-maximal ATPase activity shifts about twofold (from 4 to 2 μ M), the concentration needed for half-maximal velocity is only slightly altered. Because fewer rigid crossbridges should be needed to achieve the half-maximal velocity, the efficiency of force generation by single-headed myosin during sliding under loaded conditions may be somewhat smaller.

We thank Professors F. Oosawa and T. Mitsui and Drs E. Prochniewicz-Nakayama and A. Inoue for useful discussions. We also thank Professor Th. Wieland for phalloidin-rhodamine, and Dr R. D. Vale for critical reading of the manuscript.

Note added in proof: We recently succeeded in preparation of highly purified single-headed myosin (contaminating double-headed myosin was <0.5%). The results using such a highly purified single-headed myosin were the same as those shown here.

Received 19 January; accepted 26 February 1987.

1. Cooke, R. & Franks, K. E. *J. molec. Biol.* **120**, 361–373 (1978).
2. Tokiwa, T. & Morales, M. F. *Biochemistry* **10**, 1722–1727 (1971).

3. Margossian, S. S. & Lowey, S. J. *molec. Biol.* **74**, 312-329 (1973).
4. Gadasi, H., Oplatka, A., Lamed, R. & Muhrad, A. *Biochem. biophys. Res. Commun.* **58**, 913-918 (1974).
5. Chaen, S., Shimada, M. & Sugi, H. *J. biol. Chem.* **261**, 13632-13636 (1986).
6. Kron, S. J. & Spudich, J. A. *Proc. natn. Acad. Sci. U.S.A.* **83**, 6272-6276 (1986).
7. Honda, H., Nagashima, H. & Asakura, S. *J. molec. Biol.* **191**, 131-133 (1986).
8. Yanagida, T., Nakase, M., Nishiyama, K. & Oosawa, F. *Nature* **307**, 58-60 (1984).
9. Thames, M. D., Teichholz, L. E. & Podolsky, R. J. *J. gen. Physiol.* **63**, 509-530 (1974).
10. Yanagida, T., Arata, T. & Oosawa, F. *Nature* **316**, 366-369 (1985).
11. Higashi-Fujime, S. *J. Cell Biol.* **101**, 2335-2344 (1985).
12. Moss, R. L. & Haworth, R. A. *Biophys. J.* **45**, 733-742 (1984).
13. Maruyama, K. & Weber, A. *Biochemistry* **11**, 2990-2997 (1972).
14. Nagashima, H. *J. Biochemistry* **100**, 1023-1029 (1986).
15. Spudich, J. A. & Watt, S. J. *J. biol. Chem.* **246**, 4866-4871 (1971).
16. Huxley, H. E. *J. molec. Biol.* **7**, 281-308 (1963).
17. Eggleston, P., Elsdon, S. R. & Gough, N. *Biochem. J.* **37**, 526-529 (1943).

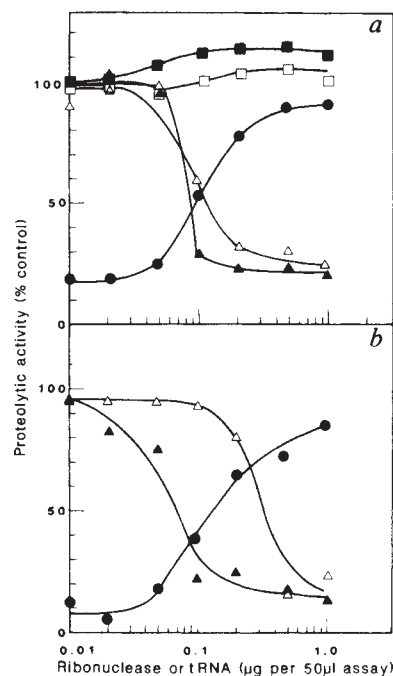


Fig. 1 Effect of ribonucleases and tRNA on the degradation of bovine and human α -LAs and of κ -IgG. *a*, Reaction mixture was preincubated for 45 min with the indicated concentrations of RNase A ($\blacktriangle$, $\blacksquare$) or micrococcal nuclease ($\triangle$, $\square$) before addition of ubiquitin and either 125 I-labelled α -LA ($\blacktriangle$, $\triangle$) or 125 I-labelled human α -LA ($\blacksquare$, $\square$). For reconstitution of the system with tRNA ($\bullet$), the reaction mixture was preincubated for 45 min with micrococcal nuclease (1 μ g), followed by inhibition of the nuclease with EGTA (or pTp) for 15 min, and initiation of the reaction with the addition of ubiquitin, 125 I-labelled bovine α -LA and tRNA (in the indicated amounts). Proteolytic activity of 100% (57% and 42% of the bovine and human substrates, respectively) were degraded) is the value that was measured in a system to which nucleases were not added; 12% and 9% of the degradation of the bovine and human α -LAs, respectively, was independent of ubiquitin and ATP. These values were subtracted from all samples and the results were normalized to 100%. *b*, Same as *a*, except that 125 I-labelled κ -IgG was the proteolytic substrate: incubations with RNase A ($\blacktriangle$), micrococcal nuclease ($\triangle$) and tRNA ($\bullet$) as above. The κ -IgG was purified to homogeneity from the urine of a patient with IgG multiple myeloma³². It was found that 39% of the substrate was degraded, and 7% of the degradation was independent of ubiquitin and ATP. Results were corrected and normalized as described for *a*.

Methods The degradation of 125 I-labelled bovine and human α -LAs was determined as described^{7,8}. Briefly, the reaction mixture contained (in a final volume of 50 μ l) 50 mM Tris-HCl pH 7.6, 1 mM CaCl_2 , 3 mM dithiothreitol (DTT), 0.5 mM ATP, 10 mM creatine phosphate (CP), 10 μ g creatine phosphokinase (CPK), 5 μ g ubiquitin, 1 μ g of 125 I-labelled substrate ($5\text{--}20 \times 10^4$ c.p.m.), and 220 μ g of fraction II protein. Following incubation at 37 $^\circ\text{C}$ for 2 h, the reaction was terminated by the addition of 0.85 ml of 15% trichloroacetic acid (TCA) in the presence of 10 mg carrier bovine serum albumin (BSA) (0.1 ml). Radioactivity was determined in a γ -counter in a 0.5 ml sample of the supernatant, after centrifugation for 5 min in a microcentrifuge.

terminal aspartic acid²³), soybean trypsin inhibitor (STI; aspartic acid²⁴) and bovine α -lactalbumin (bovine α -LA; glutamic acid²⁵). We postulated that a tRNA-dependent post-translational modification of the acidic α -NH₂ groups of proteolytic substrates is required before they are processed by the ubiquitin system.

Both bacteria and eukaryotes contain an unusual class of enzymes, aminoacyl-tRNA-protein transferases, which catalyse post-translational conjugation of specific amino-acid residues to the mature amino termini of acceptor proteins²⁶. Among the

Role of arginine-tRNA in protein degradation by the ubiquitin pathway

Sarah Ferber & Aaron Ciechanover*

Unit of Biochemistry, Faculty of Medicine, Technion-Israel Institute of Technology, PO Box 9649, Haifa 31096, Israel

Degradation of intracellular proteins through the ubiquitin and ATP-dependent proteolysis pathway involves several steps. Initially, ubiquitin is covalently linked to the proteolytic substrate in an ATP-requiring reaction. Proteins marked by ubiquitin may then be selectively lysed in a reaction that also requires ATP (for reviews see refs 1-3). A major question concerns the structural features of a protein that make it a specific substrate for ubiquitin-mediated degradation. It was shown that a free α -NH₂ group is one important feature of the protein structure recognized by the ubiquitin ligation system^{4,5}, and that the half-life *in vivo* of a protein with an exposed amino terminus depends on its amino terminal residue⁶. We have previously demonstrated that transfer RNA (tRNA) is essential for conjugation of ubiquitin and for the subsequent degradation of proteins with acidic amino termini (aspartate or glutamate)^{7,8}. We now show that tRNA is required for post-translational conjugation of arginine to acidic amino termini of proteins, a modification that is essential for their degradation by the ubiquitin pathway.

A nonlysosomal, ATP-dependent proteolytic system from rabbit reticulocytes has been recently characterized and partially purified. One of the components of the system has been identified as ubiquitin^{9,10}, an evolutionarily highly conserved polypeptide¹¹⁻¹³. Ubiquitin becomes conjugated to proteins in an ATP-requiring reaction, in which several ubiquitin molecules are bound by isopeptide linkages to ϵ -NH₂ groups of lysine residues of the protein substrate^{14,15}. Proteins thus marked by ubiquitin are selectively degraded in an ATP-dependent reaction by a specific protease^{16,17}. The relationship between ubiquitin conjugation and protein breakdown has been amply documented¹⁸⁻²⁰.

The intermediary steps in ubiquitin conjugation have been partially elucidated. In the first step, ubiquitin is activated to a high energy thiol-ester intermediate by two enzymes, E₁ and E₂ (refs 21 and 22). In the second step, a third enzyme, E₃, catalyses the formation of isopeptide bonds between the high-energy ubiquitin intermediate and the protein substrate^{5,22}. E₃ has a selective substrate binding site which is probably essential for recognition of structural features of the substrate that render it susceptible to ubiquitin-dependent proteolysis⁵.

In previous studies it was found that a free α -NH₂ group of the protein substrate is an important structural determinant for recognition by the ubiquitin system⁴. In recent studies⁸, we found that tRNA is essential for conjugation of ubiquitin, and for the subsequent degradation of three proteins with acidic amino termini, namely bovine serum albumin (BSA; with amino-

* To whom correspondence should be addressed.

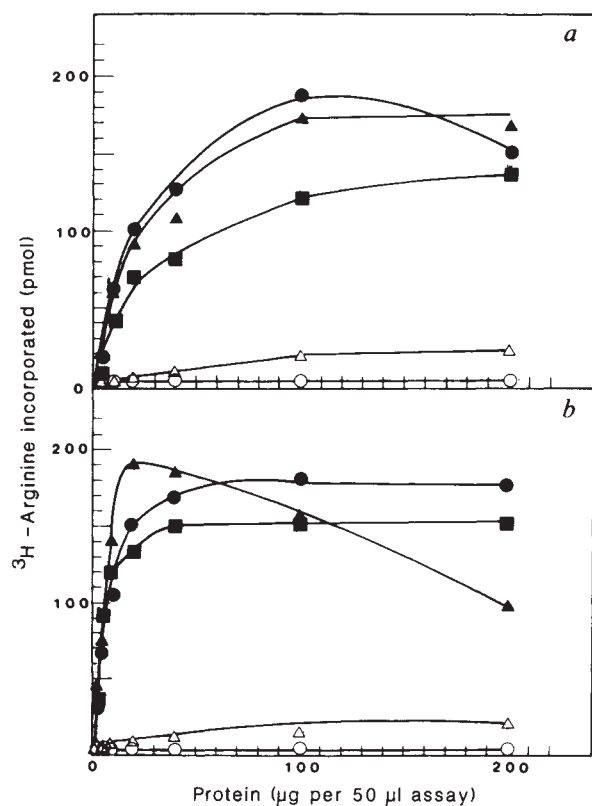


Fig. 2 Incorporation of ^3H -arginine into selective substrates of the ubiquitin pathway. *a*, Fraction II (in the indicated amounts of protein) was incubated with 330 pmol of $\text{L}[5(\text{n})\text{-}^3\text{H}]\text{arginine}$ monohydrochloride (15 Ci mmol^{-1} ; Amersham) and with 50 μg of BSA ($\bullet$), bovine α -LA ($\blacksquare$), STI ($\blacktriangle$), α -casein ($\triangle$), and either β -LG, oxidized RNase A, lysozyme, κ -casein, human α -LA, or β -casein ($\circ$). Incorporation of radioactivity into TCA-insoluble material was determined as described below. *b*, As *a*, except that fraction II (220 μg of protein) was incubated with the indicated amounts of the substrate proteins.

Methods The reaction mixture also contained (in a final volume of 50 μl) 50 mM Tris-HCl pH 7.6, 3 mM DTT, 5 mM MgCl_2 , 0.5 mM ATP, 10 μg CPK, 10 mM CP and 5 μg tRNA (calf liver; Boehringer). Following 2-h incubation at 37 $^\circ\text{C}$, 20 μl aliquots were sampled (in duplicate) onto paper disks (3MM). The disks were incubated for 5 min in 5% TCA at 90 $^\circ\text{C}$, followed by two washes in 5% TCA at 4 $^\circ\text{C}$, and a final wash in ethanol ether (50% v/v). Following drying, the disks were counted in a liquid scintillation counter. Incorporation of label into endogenous fraction II proteins (as determined in a system to which exogenous substrate was not added) did not exceed 3% of the maximal incorporation in the presence of exogenous substrates. These values were subtracted from all results, and net substrate-dependent incorporation shown.

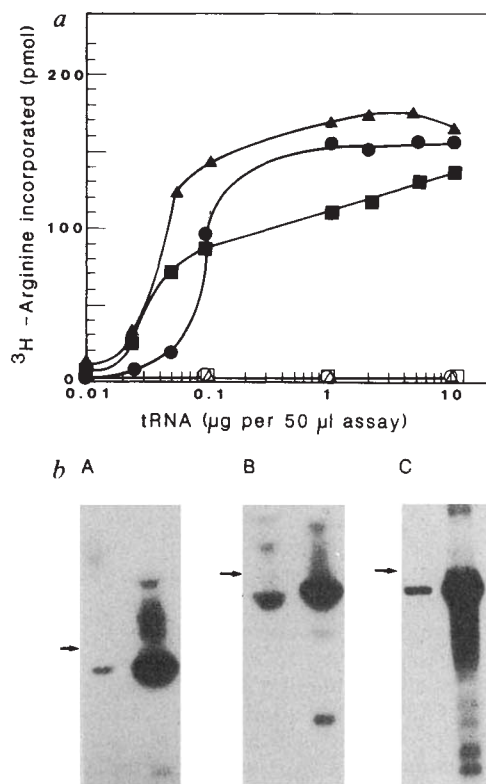


Fig. 3 The tRNA-dependent incorporation of ^3H -arginine into selective proteolytic substrates of the ubiquitin pathway. *a*, Reaction mixture preincubated for 45 min with 1 μg of MNase, followed by inhibition of the nuclease with EGTA (or pTp) for 15 min, and initiation of the reaction with the addition of 330 pmol of ^3H -arginine, 50 μg of bovine α -LA ($\blacktriangle$, $\triangle$), STI ($\blacksquare$, $\square$), or BSA ($\bullet$, $\circ$), and tRNA in the indicated amounts. ATP and an ATP-regenerating system were added to ($\blacktriangle$, $\blacksquare$, $\bullet$), or omitted from ($\triangle$, $\square$, $\circ$), the system (see Figs 1 and 2 legends). *b*, Electrophoresis of ^3H -arginylated and ^{125}I -labelled substrates. The products of the reaction containing bovine α -LA (A), STI (B), and BSA (C), were electrophoresed and fluorographed (lanes 1). Parallel ^{125}I -labelled substrates were also analysed (lanes 2).

Methods. In *a*, the reaction mixture contained 220 μg of fraction II protein and 50 μg of the substrate, and was otherwise similar to that described for Fig. 2 (except that ATP and an ATP regenerating system were omitted where indicated). Measurement of incorporation of ^3H -arginine into substrates and correction for incorporation of label into fraction II endogenous proteins were carried out as described in Fig. 2 legend. In *b*, following 2-h incubation at 37 $^\circ\text{C}$, 6 μl of 100% TCA was added to a reaction mixture similar to that described for *a* except that the system was not treated with MNase, and contained 100 μg protein of fraction II 250 μg of substrate protein and 10 μg tRNA. The mixture was incubated for 5 min at 90 $^\circ\text{C}$, chilled on ice for 30 min, and centrifuged in a microcentrifuge. Following two washes with 5% TCA, the pellet was dissolved in 100 μl 1M Tris-HCl pH 9, and the mixture was dialysed against 20 mM Tris-HCl pH 7.2, 150 mM NaCl and 0.1 mM unlabelled arginine, following by dialysis against 20 mM Tris-HCl pH 7.2. A sample containing 20,000 c.p.m. was electrophoresed on polyacrylamide gels³³ and fluorographed (using Amplify, Amersham). The ^{125}I -labelled substrates (8,000 c.p.m.) were electrophoresed in parallel lanes. Left-hand two, 15% polyacrylamide; right-hand 10% polyacrylamide.

eukaryotic enzymes, the best studied is the arginyl tRNA-protein transferase^{26,27}. We tested the hypothesis that tRNA-dependent post translational addition of an arginine moiety to an acidic terminus of a substrate, is required for its subsequent degradation through the ubiquitin system.

As can be seen in Fig. 1*a*, ubiquitin-dependent degradation of bovine α -LA is strongly inhibited by ribonuclease A (RNase A) and by micrococcal nuclease. Addition of tRNA to a system which was pretreated with micrococcal nuclease and in which the nuclease was subsequently inhibited by EGTA^{7,8}, or by thymidine-3'-5'-diphosphate (pTp)^{7,8}, almost completely restored the proteolytic activity. Human α -LA is a protein with 75% position homology to the bovine protein. They differ, however, in their amino termini. Whereas the bovine protein has glutamic acid in this position²⁵, lysine is the amino terminus of the human protein²⁸. As can be seen in Fig. 1*a*, the ubiquitin-dependent degradation of human α -LA is not inhibited by ribonucleases.

For further confirmation of our hypothesis that proteins with acidic amino termini require tRNA for their degradation, we searched for additional substrates with similar α -NH₂ groups. One such protein is the κ -light chain of the human immunoglobulin molecule (κ -IgG) which has either glutamic acid or aspartic acid in this position²⁹. The ubiquitin- and ATP-dependent degradation of κ -IgG was inhibited by ribonucleases and could be restored by the addition of tRNA to a

Table 1 Effect of micrococcal nuclease on the degradation of ^{125}I -labelled and ^3H -arginine-labelled substrates of the ubiquitin pathway

System	Substrate Reaction conditions	$[^{125}\text{I}]\alpha\text{-LA}$	$^3\text{H}\text{-Arg } \alpha\text{-LA}$	$[^{125}\text{I}]\text{STI}$	$^3\text{H}\text{-Arg-STI}$	$[^{125}\text{I}]\text{BSA}$	$^3\text{H}\text{-Arg-BSA}$
1	Complete system	68	74	26	50	22	40
2	-ubiquitin	19	16	7	23	3	9
3	-ATP	16	13	5	20	1	4
4	+MNase + EGTA (or pTp)	22	79	8	53	2	44
5	+MNase + EGTA (or pTp) + tRNA	67	—	20	—	15	—

MNase, micrococcal nuclease. Values given are degradation in per cent. Degradation was determined in a reaction mixture similar to that described in Fig. 1 legend (the complete system, system 1). In systems 2 and 3, ubiquitin or ATP and an ATP-regenerating system were omitted, respectively. In system 4, the reaction mixture was preincubated for 45 min with 1 μg MNase before inhibition of the nuclease (see Fig. 1 legend), and addition of ubiquitin and the labelled substrate. System 5: as 4, except that 5 μg tRNA was added, together with ubiquitin and the labelled substrate. **Methods.** The ^3H -arginylated substrates were prepared as described in Fig. 3. The reaction mixture contained 5 μg ($2\text{--}4 \times 10^4$ c.p.m.) of the tritiated substrates. The ^{125}I -labelled substrates were treated in a similar way, except that tRNA was omitted from the reaction mixture, and fraction II was preincubated for 30 min with 5 μg of MNase before addition of the labelled substrate to prevent arginylation of the ^{125}I -labelled substrates.

system in which the nuclease activity had previously been inhibited (Fig. 1b).

We next attempted to demonstrate a substrate-specific amino-terminal modifying activity in crude reticulocyte fraction II (which contains all the ubiquitin system components except for ubiquitin^{1,7,8}). Fraction II does indeed contain an arginyl tRNA-protein transferase activity which is specific to substrates with acidic amino termini (Fig. 2). The arginylation was dependent on enzyme and substrate concentration (Fig. 2), metabolic energy and tRNA (Fig. 3a). The incorporation of arginine was almost completely dependent on the addition of the appropriate exogenous substrate; very little incorporation into fraction II endogenous proteins was detected (see Fig. 2 legend). Electrophoretic analysis of the reaction products demonstrated almost exclusive incorporation of the label into the exogenously added substrate (Fig. 3b). A single cycle of manual Edman degradation³⁰ showed that >89% of the arginine label was incorporated into the amino-terminal position of the various proteins (data not shown). Similar results were obtained using $\kappa\text{-IgG}$ as a substrate (S. Elias and A.C., unpublished data).

Next we examined whether tRNA-dependent arginylation of substrates with acidic amino termini is required for their degradation by the ubiquitin system and whether ribonucleases inhibit the degradation of these substrates by destroying the tRNA component of the system and so preventing this initial modification. The degradation of ^{125}I -labelled BSA, STI and bovine $\alpha\text{-LA}$ was sensitive to ribonuclease and could be restored by addition of exogenous tRNA (Table 1). In striking contrast, the degradation of the same substrates in which the amino termini were modified by the addition of ^3H -arginine was insensitive to ribonucleases.

We have shown that tRNA-dependent addition of arginine to acidic amino termini of proteins is required for their degradation by the ubiquitin system. Such a modification may be necessary for recognition and binding of the substrate to ubiquitin-protein ligase (E_3) before conjugation with ubiquitin. It is possible that the site of E_3 to which the amino terminus of the protein substrate binds is negatively charged, so that only proteins with positively charged amino termini can bind to it. Proteins with neutral or basic amino terminal residues fulfil this requirement. Proteins with acidic amino-terminal residues have no net charge at this site, and therefore cannot bind to E_3 . Indeed, a selective binding site of E_3 for proteins with basic amino termini has recently been identified (Y. Reiss, D. Kaim and A. Hershko, unpublished results).

It is not clear yet whether other amino acids can participate in post-translational modification of substrates of the ubiquitin system. We have previously reported that tRNA^{His} is essential for degradation of BSA in the ubiquitin system⁷. In preliminary studies we have found a tRNA- and energy-dependent histidyla-

tion of substrates with acidic amino termini (S. Arfin and A.C., unpublished results). The relevance of this modification to ubiquitin-dependent proteolysis is being investigated. Modification of proteins by lysine and leucine has also been reported³¹, although its relevance to proteolysis by the ubiquitin system is unknown.

Recently, it has been reported that the half-life *in vivo* of a protein is a function of its amino-terminal residue. An amino-end rule was proposed, according to which long-lived proteins have 'stabilizing' amino termini, whereas short-lived proteins have 'destabilizing' amino-terminal residues⁶. Genetically engineered β -galactosidase with glutamic acid or aspartic acid in its $\alpha\text{-NH}_2$ position had a short half-life (10–30 min)⁶. It is possible that aspartic or glutamic acid are not 'destabilizing' residues as such, but become so only through their ability to be conjugated to other 'destabilizing' residues. Note that β -galactosidase with arginine at its amino terminus had a half-life of <2 min. Caution is necessary with the interpretation that the amino-terminal residue of a protein is the only recognition signal for ubiquitin ligation⁶. For example, RNase A (which has a lysine, a strong destabilizing residue in the $\alpha\text{-NH}_2$ position) does not bind to E_3 and is not degraded by the ubiquitin system. Oxidation of methionine residues causes the protein to bind tightly to E_3 and to become an efficient proteolytic substrate of the system⁵.

We thank Dr Avram Hershko for his advice Dr Stuart Arfin for his help and for reading the manuscript, Dr Robert Fallon (Washington University School of Medicine) for his help obtaining the urine from the multiple myeloma patient, Mrs Deborah Ganoh for purifying $\kappa\text{-IgG}$ and Mrs Atara Katznelson for typing the manuscript. The research was supported by a grant from the United States-Israel Binational Science Foundation and by grants from the Vice President of the Technion for Research (Lovengart Foundation), and the Foundation for Promotion of Research in the Technion. A.C. is a Research Career Development Awardee of the Israel Cancer Research Fund, USA.

Received 26 November 1986; accepted 20 February 1987.

1. Hershko, A. & Ciechanover, A. *A Rev. Biochem.* **51**, 335–364 (1982).
2. Ciechanover, A., Finley, D. & Varshavsky, A. *J. cell. Biochem.* **24**, 27–53 (1984).
3. Hershko, A. & Ciechanover, A. *Prog. molec. Biol. Nucleic Acid Res.* **33**, 19–56 (1986).
4. Hershko, A., Heller, H., Eytan, E., Kaklij, G. & Rose, I. A. *Proc. natn. Acad. Sci. U.S.A.* **81**, 7021–7025 (1984).
5. Hershko, A., Heller, H., Eytan, E. & Reiss, Y. *J. biol. Chem.* **261**, 11992–11999 (1986).
6. Bachmair, A., Finley, D. & Varshavsky, A. *Science* **234**, 179–186 (1986).
7. Ciechanover, A., Wolin, S. L., Steitz, J. A. & Lodish, H. F. *Proc. natn. Acad. Sci. U.S.A.* **82**, 1341–1345 (1985).
8. Ferber, S. & Ciechanover, A. *J. biol. Chem.* **261**, 3128–3134 (1986).
9. Ciechanover, A., Elias, S., Heller, H., Ferber, S. & Hershko, A. *J. biol. Chem.* **255**, 7525–7528 (1980).
10. Wilkinson, K. D., Urban, M. K. & Haas, A. L. *J. biol. Chem.* **255**, 7529–7532 (1980).
11. Goldstein, G. *et al. Proc. natn. Acad. Sci. U.S.A.* **72**, 11–15 (1975).
12. Özkaynak, E., Finley, D. & Varshavsky, A. *Nature* **312**, 663–666 (1984).
13. Wiborg, O. *et al. EMBO J.* **4**, 755–759 (1985).

14. Ciechanover, A., Heller, H., Elias, S., Haas, A. L. & Hershko, A. *Proc. natn. Acad. Sci. U.S.A.* **77**, 1365-1368 (1980).
15. Hershko, A., Ciechanover, A., Heller, H., Haas, A. L. & Rose, I. A. *Proc. natn. Acad. Sci. U.S.A.* **77**, 1783-1786 (1980).
16. Hershko, A., Leshinsky, E., Ganoth, D. & Heller, H. *Proc. natn. Acad. Sci. U.S.A.* **81**, 1619-1623 (1984).
17. Hough, R., Pratt, G. & Rechsteiner, M. *J. biol. Chem.* **261**, 2400-2408 (1986).
18. Hershko, A., Eytan, E., Ciechanover, A. & Haas, A. L. *J. biol. Chem.* **257**, 13964-13970 (1982).
19. Chin, D. T., Kuehl, L. & Rechsteiner, M. *Proc. natn. Acad. Sci. U.S.A.* **79**, 5857-5861 (1982).
20. Ciechanover, A., Finley, D. & Varshavsky, A. *Cell* **37**, 57-66 (1984).
21. Ciechanover, A., Elias, S., Heller, H. & Hershko, A. *J. biol. Chem.* **257**, 2537-2542 (1982).
22. Hershko, A., Heller, H., Elias, S. & Ciechanover, A. *J. biol. Chem.* **258**, 8206-8214 (1983).
23. Shearer, W. T., Bradshaw, R. A., Gurd, F. R. N. & Peters, T. *J. biol. Chem.* **242**, 5451-5459 (1967).
24. Odani, S. & Ikenaka, S. *J. Biochem.* **71**, 839-848 (1972).
25. Brew, K., Castellino, F. J., Vanaman, T. C. & Hill, R. L. *J. biol. Chem.* **245**, 4570-4582 (1970).
26. Soffer, R. L. in *Transfer RNA: Biological Aspects* (eds Soll, D., Abelson, J. N. & Schimmel, P.) 493-505 (Cold Spring Harbor Laboratory, New York, 1980).
27. Kaji, H. *Biochemistry* **15**, 5121-5125 (1976).
28. Findlay, J. B. C. & Brew, K. *Eur. J. Biochem.* **27**, 65-86 (1972).
29. Kabat, E. A., Wu, T. T., Bilofsky, H., Reid-Miller, M. & Perry, H. in *Sequences of Proteins of Immunological Interest*, U.S. DHHS, PHS. 14-29. National Institute of Health (1983).
30. Glazer, A. N., DeLange, R. J. & Sigman, D. S. in *Chemical Modification of Proteins* (eds Work, T. S. & Work, E.) 60-64 (North Holland, Amsterdam, 1975).
31. Shyne-Athwal, S., Riccio, R. V., Chakraborty, G. & Ingoglia, N. A. *Science* **231**, 603-605 (1986).
32. Hudson, L. & Hay, F. C. (eds) *Practical Immunology* 157-165 (Blackwell, Oxford, 1980).
33. Laemmli, U. K. *Nature* **227**, 680-685 (1970).

Conversion of allosteric inhibition to activation in phosphofructokinase by protein engineering

Frankie Tat-Kwong Lau & Alan R. Fersht

Department of Chemistry, Imperial College of Science and Technology, London SW7 2AY, UK

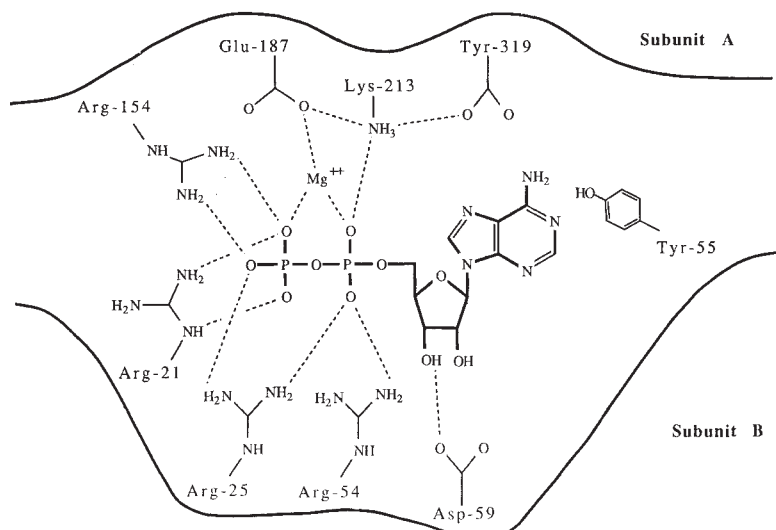
Many enzymes are subject to allosteric control, often with inhibitors and activators binding to the same effector site. Phosphofructokinase in *Escherichia coli* is such an enzyme, being inhibited by phosphoenolpyruvate (PEP) and activated by ADP and GDP¹. How do individual interactions with effectors affect the balance between activation and inhibition, especially when both ligands share aspects of the same binding site? We find that mutation of a single residue in the effector site, Glu → Ala 187, leads to PEP being an activator rather than an inhibitor. With low concentrations of the substrate fructose-6-phosphate, the mutant enzyme is more than one hundred times more active than wild-type enzyme at millimolar concentrations of PEP. The classical Monod-Wyman-Changeux two-state model² is too simple to account for the properties of the mutant enzyme.

Phosphofructokinase is a key regulatory enzyme in glycolysis. The tetrameric enzyme in *E. coli* is subject to feedback control: the binding of fructose-6-phosphate (F6P) is strongly cooperative and is allosterically inhibited by phosphoenolpyruvate (PEP) and activated by ADP or GDP. Both inhibitor and activators bind to the same effector site. Early detailed studies have shown that the enzyme conforms to a classical Monod-Wyman-Changeux *K*-system in which F6P binds poorly to the *T*-state (the predominant form of the free enzyme) but well to the *R*-state³. Both states have the same values of rate constants (k_{cat}) for bound substrates. There are advanced X-ray crystallographic studies on both the *R*- and *T*-states of the enzymes from *E. coli* and *Bacillus stearothermophilus* by Evans and colleagues^{4,5}. The crystal structure of the enzyme from *E. coli* with bound ADP shows that the effector site lies between two subunits and that ADP makes extensive interactions with both (Fig. 1 and P. R. Evans, unpublished data). Little is known about the precise mode of binding of PEP to the effector site. The PEP analogue phosphoglycolate has been diffused into *R*-state crystals of the *Bacillus stearothermophilus* phosphofructokinase and the structure of the complex solved at 7 Å resolution, which is too low for fine details⁴. The analogue appears to bind to the same site as ADP, and model building suggests the phosphate group of PEP may bind to the same site as the β -phosphate group of ADP⁵.

Studies by Perutz and co-workers have pointed the way to understanding allosteric interactions. Using natural point mutants and products of site-directed mutagenesis, they have found residues with key functions in cooperativity⁶. Similarly, we have begun systematically mutating the effector site of phosphofructokinase to delineate the structural factors that balance activation and inhibition. The target residues in this study are Arg 21, Arg 25, Lys 213 and Glu 187. We have mutated each in turn to an alanine residue to remove the charged side chain. The three positively charged residues are involved in binding the α - and β -phosphate groups of ADP (Fig. 1) and, by analogy, also bind those in GDP. (GDP is used in preference to ADP in kinetic studies of the effector site because of product inhibition by ADP.) The negatively charged side chain of Glu 187 binds the magnesium ion that is coordinated to the nucleotide.

Mutation of all four residues gives active enzyme. Truncation of Lys 213 renders the binding of GDP and PEP undetectable up to 3 mM. The loss of Arg 25 weakens the binding of GDP 30-fold and renders the binding of PEP undetectable up to 20 mM. Mutation of Arg 21 abolishes the binding of GDP and weakens that of PEP. The mutant Glu → Ala 187 has more interesting properties (Table 1): the binding of GDP to the mutant is not detectable at millimolar concentrations, either from direct

Fig. 1 Simplified representation of the effector site of *E. coli* phosphofructokinase with bound ADP (P. R. Evans, personal communication). ADP binds between two subunits as indicated. The Mg^{2+} ion is octahedrally coordinated and makes additional contacts with two water molecules and the backbone carbonyl group of Gly 185 (not shown).



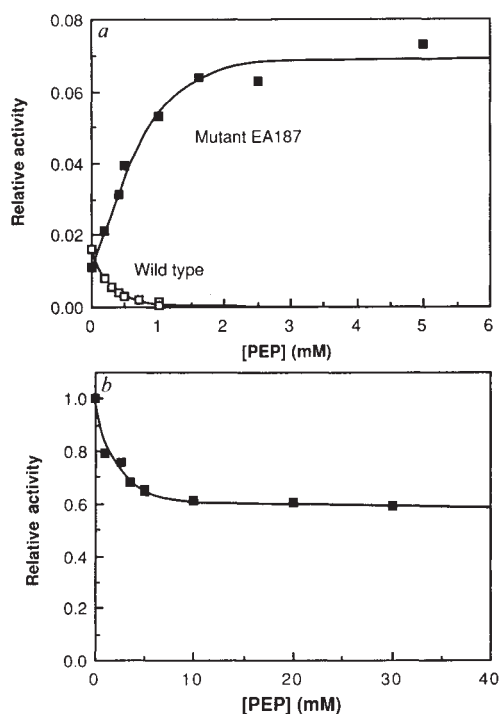


Fig. 2 *a*, Dependence on PEP concentrations of the activity of wild-type ($\square$), and mutant Glu $\rightarrow$ Ala 187 ($\blacksquare$), phosphofructokinases in the presence of 0.3 mM F6P and 1 mM ATP, at pH 8.2 and 25°C. *b*, As *a*, but with 1 mM F6P (close to saturation). Activity of wild-type enzyme is very low above 1 mM PEP.

measurements or from competition with the binding of PEP (see below), even though the dissociation constant ($K_{R(GDP)}$) from wild-type enzyme is 40 μ M. Mutation of Glu $\rightarrow$ Ala 187 has a remarkable effect on the consequences of binding PEP. At low concentrations of F6P, as found *in vivo*, where the enzyme is predominantly in the *T*-state, PEP is an activator. At 0.3 mM PEP, for example, there is a sixfold increase in activity on binding PEP (Fig. 2). Wild-type enzyme is strongly inhibited under the same conditions (Fig. 2). At high concentrations of F6P, where the mutant enzyme is close to being saturated with F6P and is predominantly in the *R*-state, there is some inhibition as a result of the binding of PEP (with a 40% reduction in rate at saturating concentrations of PEP). This is quite different from the inhibition of wild-type enzyme where the activity tends to zero.

Mutation of Glu $\rightarrow$ Ala 187 affects homotropic interactions of binding of F6P. In the absence of PEP there is a slight decrease in the Hill coefficient (n_H) from 3.7 in wild-type enzyme to 3.0 in the mutant. There is a progressive decrease in the cooperativity of F6P binding to mutant as the concentration of PEP increases, falling to only 1.1 at 10 mM PEP. The concentration of F6P required for half maximal activity ($S_{1/2}$) decreases slightly with increasing PEP concentrations. This behaviour contrasts with the effects of PEP on wild-type enzyme for which $S_{1/2}$ increases greatly as the concentration of PEP is raised.

According to the classical analysis of allostery by Monod, Wyman and Changeux², PEP binds to the *T*-state of wild-type enzyme^{2,3}. But the effects of PEP on the mutant are consistent with PEP binding to an *R*-state. The data are not consistent, however, with the conventional two-state allosteric model² for the mutant because the *R*-state with PEP bound has a lower value of k_{cat} than does the *R*-state with just F6P bound. One could invoke the existence of more quaternary states, for example *T*, *R* and *R'* in a three-state model. But it is simpler to consider that the enzyme exists in two basic quaternary structures whose precise kinetic and thermodynamic properties vary slightly according to the reaction conditions and ligands present.

Table 1 Kinetic properties of wild-type and mutant enzymes

Enzyme	Specific activity* (μ mol $\text{min}^{-1} \text{mg}^{-1}$)	Hill constant, n_H	$S_{1/2}(\text{F6P})^\dagger$ (mM)
Wild-type	190	3.7	0.33
Glu $\rightarrow$ Ala 187	87	3.0	0.75
Glu $\rightarrow$ Ala 187 + 10 mM PEP		1.1	0.57

* Production of fructose-1,6-bisphosphate in the presence of 1 mM F6P, 1 mM ATP at 25°C and pH 8.2 (ref. 8).

† Substrate concentration for half V_{max} .

Methods. The gene *pfkA* (ref. 9) was recloned in pEMBL8(+)¹⁰ to give pHL1 which was constructed to give high level, inducible, expression of phosphofructokinase (data not shown). For mutagenesis, a synthetic primer which contained a single mismatch with the wild-type gene was designed to introduce an Ala codon at position 187. Site-directed mutagenesis was performed as described¹¹ using TG2¹² as the recipient strain in transformation. Phosphofructokinase was expressed in HE1 cells (*pro-82*, Δ *pfkB201*, *recA56*, Δ (*rha-pfkA*)200, *endA1*, *hsdR17*, *supE44/F'*:*tra D36*, *proAB*⁺, *lacI*^q, *lacZ* Δ M15) which is a phosphofructokinase-less strain. The enzyme was purified from 24-h culture grown in 2 $\times$ TY broth containing 100 μ g ml⁻¹ ampicillin and 70 μ g ml⁻¹ isopropyl thiogalactoside (IPTG) by a Blue-A column (Amicon Matrex) and by Sephacryl S-300 gel filtration chromatography. The activity of the enzyme was assayed by a coupled enzyme method⁸.

Resolving the structural basis of the allosteric effects requires further high-resolution X-ray crystallographic studies. The mutagenesis studies have, however, provided definitive evidence that both inhibitor and activator bind in the same site. Mutation of Glu $\rightarrow$ Ala 187 has pinpointed one of the important residues and suggests that the crystal structure of the mutant is a candidate for further study. It may well reveal further conformational states that are important in function but not so accessible in wild-type enzyme. One of the goals of industrial protein engineering is to alter allosteric properties of regulated enzyme in order to alter yields of products⁷. This study shows that it is feasible to do this by systematically mutating a relatively small number of amino acids in effector binding sites.

We thank Dr P. R. Evans for his encouragement and unpublished data. F.T.-K.L. is a recipient of a scholarship from the Croucher Foundation.

Received 26 January; accepted 20 February 1987.

1. Uyeda, K. *Adv. Enzym.* **48**, 194-244 (1979).
2. Monod, J., Wyman, J. & Changeux, J. P. *J. molec. Biol.* **12**, 88-118 (1965).
3. Blangy, D., Buc, H. & Monod, J. *J. molec. Biol.* **31**, 13-35 (1968).
4. Evans, P. R., Farrants, G. W. & Lawrence, M. C. *J. molec. Biol.* **191**, 713-720 (1986).
5. Evans, P. R. & Hudson, P. J. *Nature* **279**, 500-504 (1979).
6. Nagai, K., Perutz, M. F. & Poyart, C. *Proc. natn. Acad. Sci. U.S.A.* **82**, 7252-7255 (1985).
7. Ulmer, K. M. *Science* **219**, 666-670 (1983).
8. Kotlarz, D. & Buc, H. *Meth. Enzym.* **90**, 60-70 (1982).
9. Hellinga, H. W. & Evans, P. R. *Eur. J. Biochem.* **149**, 363-373 (1985).
10. Dente, L., Cesareni, G. & Cortese, R. *Nucleic Acids Res.* **11**, 1645-1655 (1983).
11. Zoller, M. J. & Smith, M. *Meth. Enzym.* **100**, 468-500 (1983).
12. Gibson, T. J. thesis, Univ. Cambridge (1984).

Erratum

Need for DNA topoisomerase activity as a swivel for DNA replication and for transcription of ribosomal RNA

Steven J. Brill, Stephen DiNardo, Karen Voelkel-Meiman & Rolf Sternglanz

Nature **326**, 414-416 (1987).

THIS letter was published with an incorrect title. The above title is correct.